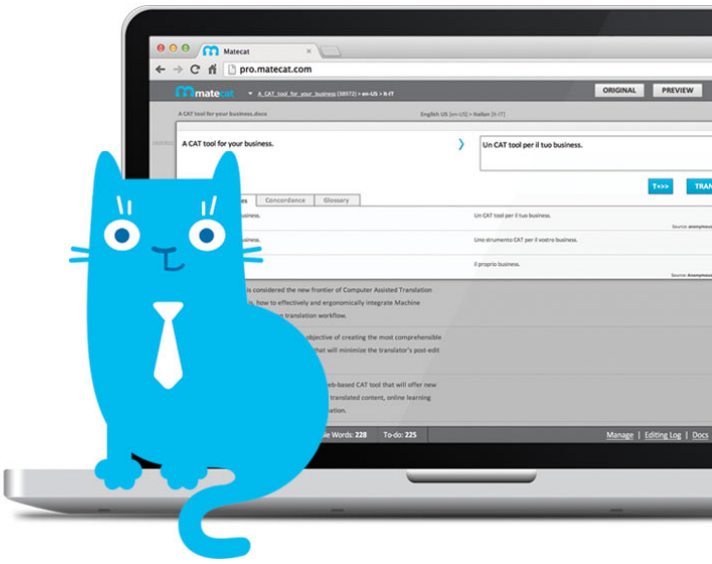
MateCat

Post-editing and outsourcing made easy



User manual and installation guide

Introducing MateCat **4**
How MateCat Calculates Payable Words 4
Volume Analysis Page 5
Supported browser, languages and formats 6
Translation Process 9
Start working with MateCat **11**
MateCat Home Page and Login 11
Creating Projects 11
How To Upload File(s) for Translation 12
Translation Memory and Machine Translation 14
Machine Translation Servers 15
Analysis Report 16
Public Translation Memory and Translation Memory Key 20
TM backup and updates 21
How to add a Glossary 22
Translation Editor User Interface 23
- Top bar 24
- Editor window 26
- Bottom bar 28

Translation Toolbox **31**
Manage your language resources 31
Working Offline 35
Using TM matches, MT suggestions and Glossary terms 36
How to Use Glossary Terms in Your Translated Segment 39
Working with Text and Tags 40
- Tag autocompletion 40

- Formatting features 40
- Adding line breaks and hidden text 41
- Placeables and untranslatable text 41

Confirming a Segment 42

Autopropagation 42

Concordance Search 43

Navigating through Files 44

QA Messages 44

- Tag Mismatch 45
- Translation conflicts 46
- Tag Order Mismatch 46
- Whitespaces mismatch 47

Revision in MateCat 48

- Revision process 48
- Quality Report 49
- Finalizing 51

Spell-checking 51

Editing Log 53

Finalising a Project 56

Managing Projects with MateCat 58

Management Panel 58

Splitting Jobs 61

Resources for LSPs 66

Appendix 68

Keyboard Shortcuts 68

Installing an open source version of MateCat 71

Introducing MateCat

MateCat is an open source, enterprise-level, web-based translation software which closely integrates machine translation and human translation. Using MateCat, you obtain between 10% and 20% more matches for your translations thanks to the largest public translation memory in the world and the best machine translation currently in use.

How MateCat Calculates Payable Words

MateCat counts words by using a **combination** of advanced **TM technology** (MyMemory) and application of a reduction in weighting **for machine translation suggestions**. This allows MateCat to give you **more matches** than any other translation tool.

According to industry standards, words or phrases with a 100% Translation Memory match are given a weighting of 30% and words or phrases with a partial TM match are given a weighting of 60%.

Here is an example to calculate weighted words[1]:

No TM match =	100%
Lower fuzzy ranges =	50 and 70%
Higher fuzzy ranges =	40%
100% match =	30%
Repetition (same segment repeated in the document) =	30%

[1] This is an example. Each CAT tool counts words differently and each LSP calculates weighted words differently.

Context match (more than 100% match, because of same context information in the TM) = 0%

MateCat counts matches according to these standards.

For Machine Translation, MateCat decides which reduction in weighting to apply on the basis of the extent to which the **MT** has been useful for the past 1 million words for each language combination. MateCat assumes that the less the machine translation suggestion is edited by translators, the more useful it is.

We decided to **split the benefits of this technology** between the language service provider and the translator. So, if a translator saves 20% of his/her time, the word count is reduced by 10% only.

How MateCat counts **payable words** in a project is indicated in the volume analysis report generated during creation of the project itself.

Volume Analysis Page

The volume analysis page which opens during the creation of a project in MateCat shows you a global **report of the volume** to be translated:

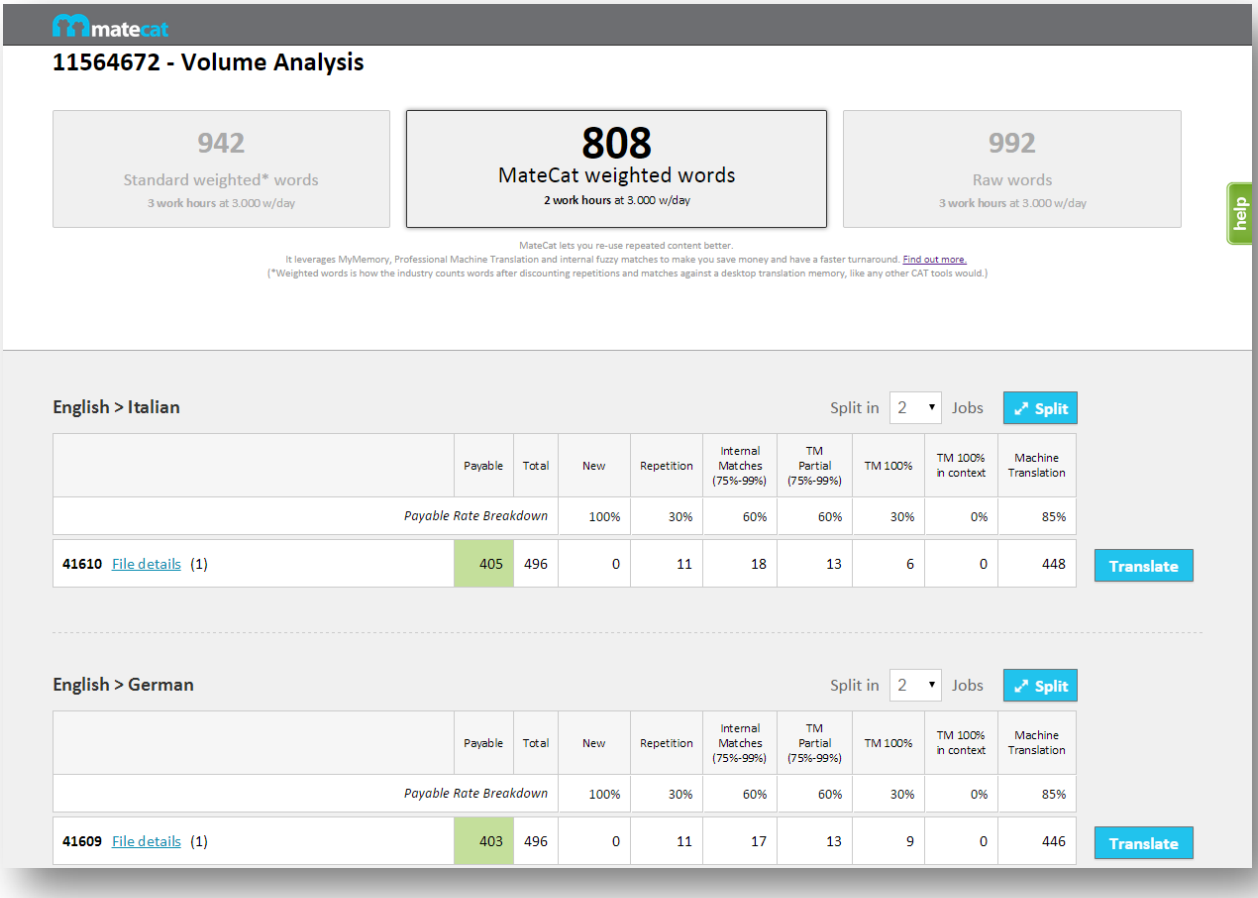
Standard weighted words: The volume as calculated by industry standard CAT tools. Calculation of the weighted word count takes into consideration repetitions and TM matches.

MateCat weighted words: The MateCat word count which not only counts repetitions and TM matches, but also generates reductions in weightings for internal matches and MT suggestions.

Raw words: The word count without any reduction in weighting for repetitions or matches.

For each language pair, you will find the details of the analysis.

In case of **multiple language pairs**, this global report shows you the total project volume.


matecat
11564672 - Volume Analysis
942
Standard weighted* words
3 work hours at 3.000 w/day
808
MateCat weighted words
2 work hours at 3.000 w/day
992
Raw words
3 work hours at 3.000 w/day
help
MateCat lets you re-use repeated content better.
It leverages MyMemory, Professional Machine Translation and internal fuzzy matches to make you save money and have a faster turnaround. Find out more.
(*Weighted words is how the industry counts words after discounting repetitions and matches against a desktop translation memory, like any other CAT tools would.)
English > Italian
Split in 2 Jobs Split
Payable Total New Repetition Internal Matches (75%-99%) TM Partial (75%-99%) TM 100% TM 100% in context Machine Translation
Payable Rate Breakdown 100% 30% 60% 60% 30% 0% 85%
41610 File details (1) 405 496 0 11 18 13 6 0 448 Translate
English > German
Split in 2 Jobs Split
Payable Total New Repetition Internal Matches (75%-99%) TM Partial (75%-99%) TM 100% TM 100% in context Machine Translation
Payable Rate Breakdown 100% 30% 60% 60% 30% 0% 85%
41609 File details (1) 403 496 0 11 17 13 9 0 446 Translate

Supported browser, languages and formats

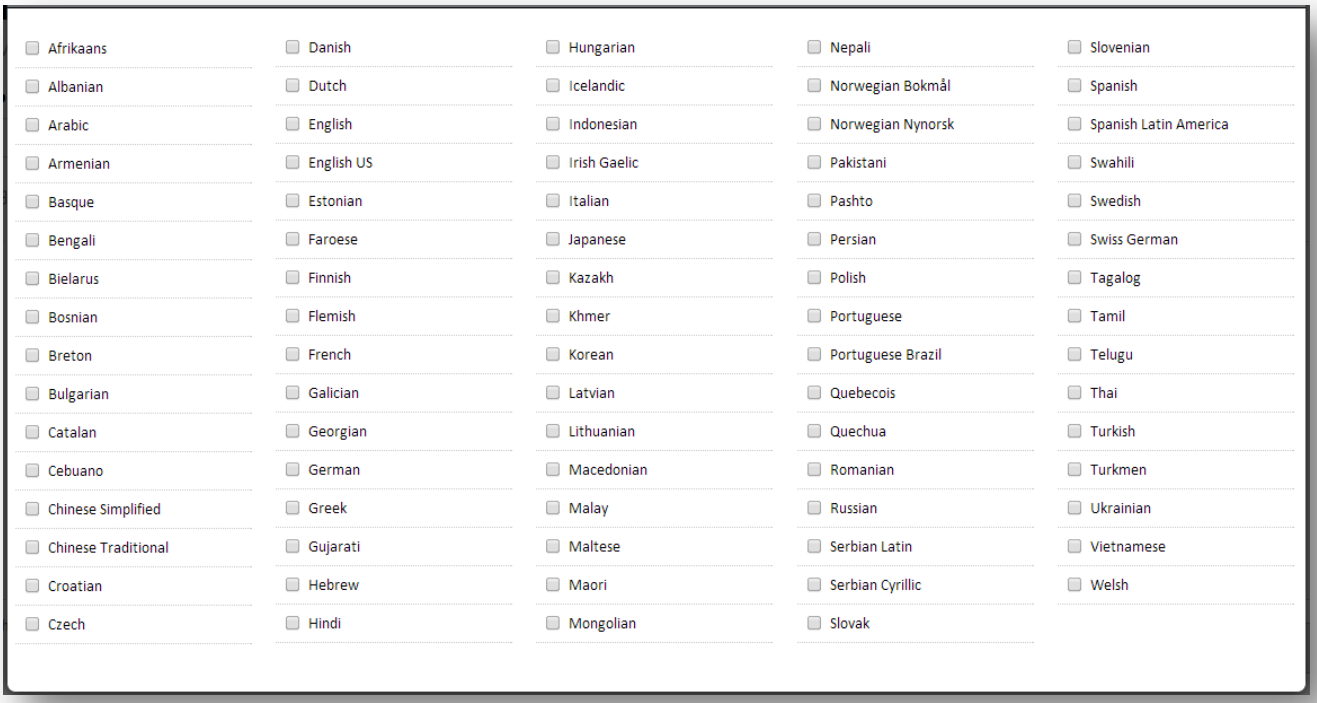
All you need to use to start working with MateCat is an up-to-date version of a supported web browser:

- **Google Chrome** 30, or higher (on Windows, Linux and Mac OS X)

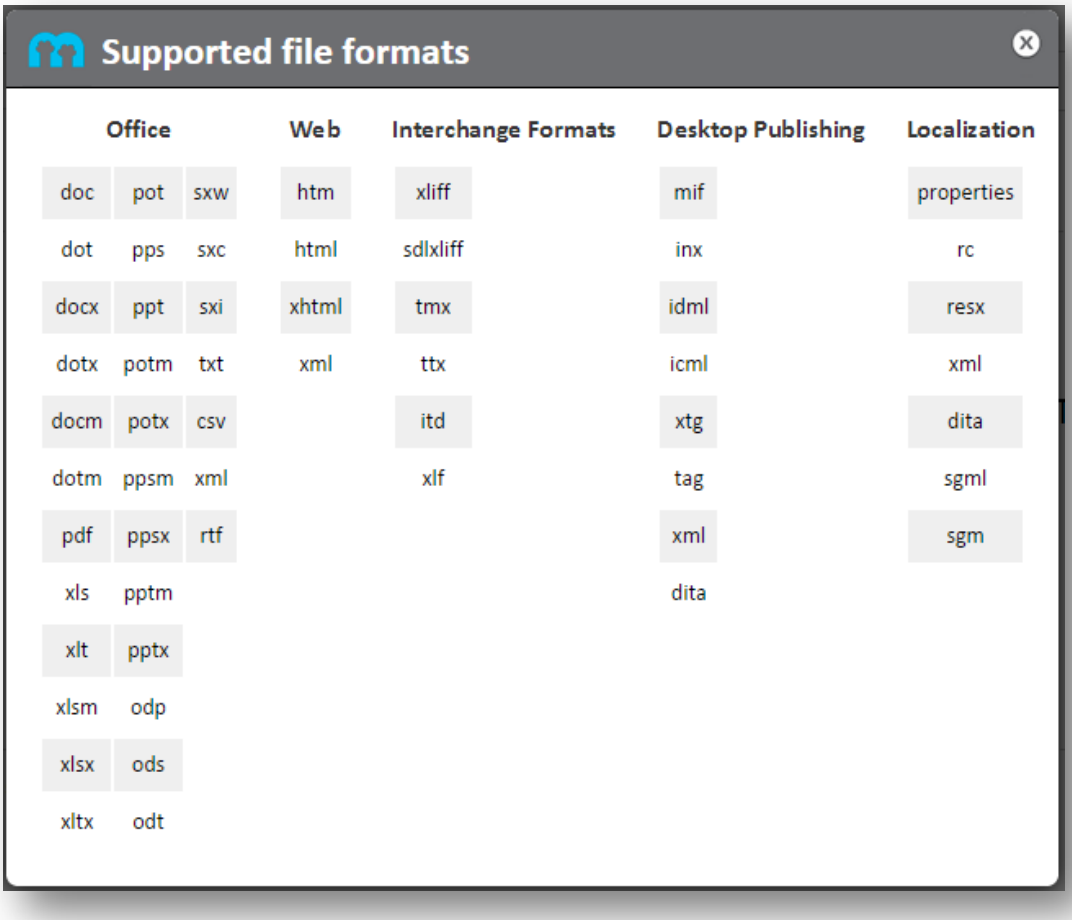
- **Safari** 5.1.7, or higher (on Mac OS X)

MateCat supports **79 languages,** works with many language combinations, and supports **57 file formats**.

Here is the list of supported languages and file formats:



Supported languages



Supported file formats

Translation Process

The translation process is divided into the following phases:

1. **Creating a project**: from the MateCat home page (http://www.matecat.com) you can select the **source and target language(s)** and add the file(s) to translate. Then you can choose one of the following options:

   - use a **machine translation server** of your choice. This option allows you to receive **machine translation suggestions** for each segment from the server you choose. You can disable this option by selecting **NONE** from the dropdown menu.
   - associate an existing or new **private TM**. By associating a private TM, you will be able to store all the translated segments in this private translation memory and add new glossary items if necessary. In this way, you will be able use this private TM for future projects and take advantage of already translated segments.
   - do **without any private TM**. In this way, you will receive suggestions and store your translation in a **public TM exclusively for MateCat users**.

If you choose to receive suggestions from both the MT and the TM, you can receive from 10 to 20% more matches for all your translations. This can increase your productivity and enable you to work more quickly.

2. **Translating**: the **Translation Editor** window shows the translation interface. It has 2 columns: source text in the left column and translated text in the right column. When you open a segment during your translation, MateCat searches for **matches from translation memories**[2] **and from machine translations**, if

[2] Both your private TM, if any, and the public shared translation memory.

enabled. During the translation, MateCat also automatically performs some **quality checks** on Tag mismatches, Tag order mismatches, Whitespace mismatches and translation conflicts.

3. **Finalising a project**: A progress bar at the bottom of the Translation Editor window shows the progress of your translation. The translation is finished when the progress bar is at **100% complete**, when there are no *To-do* words left and when **no errors are reported in the QA module** . When the translation is complete and all issues have been fixed, you can click on the **Download translation** button to download the translated file in the original format.

Start working with MateCat

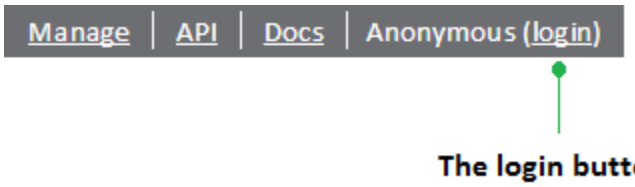
MateCat Home Page and Login

There are currently two versions of the MateCat translation tool:

- http://demo.matecat.com – Open source version to translate XLIFF files. The user can also install and administer the open source version on his/her own server. For further instructions and details, refer to the installation guide.

- http://www.matecat.com – Proprietary version which includes filters for **57 file formats**. **This is the version we use as a reference for this user manual.** This version of MateCat is free for both personal and business translation jobs.

In order to store and recover your projects in MateCat, you should **log in** by clicking on the **login** button at the bottom of the MateCat home page. You will be asked to sign in using your **Google account**.



Creating Projects

In MateCat, translation jobs are organised into **projects**. A project is made up of one or more translation **jobs** and each one is associated with a unique **job URL.** Each project can be created with one source language and **one or more target languages**.

To create a new project in MateCat, visit http://www.matecat.com

All you need to do is select the language combination and upload the file(s) to translate. This is necessary to create your project.

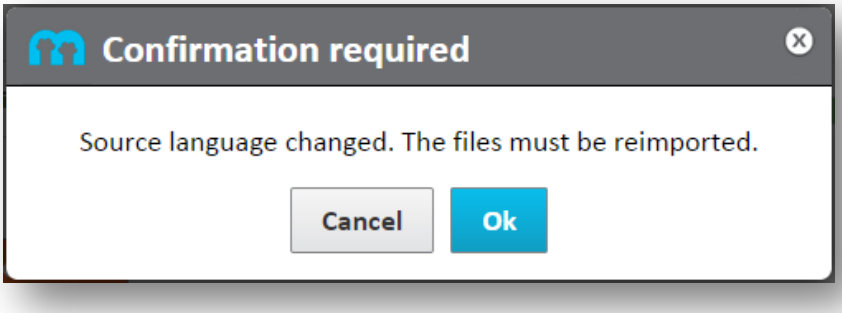
By clicking on **Options**, you can add options or change preferential options.

If you need to store your project and recover it later, you should firstly **log in** using the **Login** button at the bottom of the page. This allows you to add all your projects to your Management Panel, recover all information about projects, such as the volume analysis or job URL, and check the progress of translations.

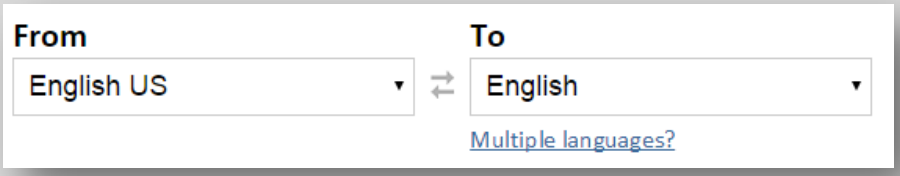
Note: it is not possible to work on MateCat offline. You can only save your translation and receive suggestions from the MT(s) and TM(s) if you work online.

How To Upload File(s) for Translation

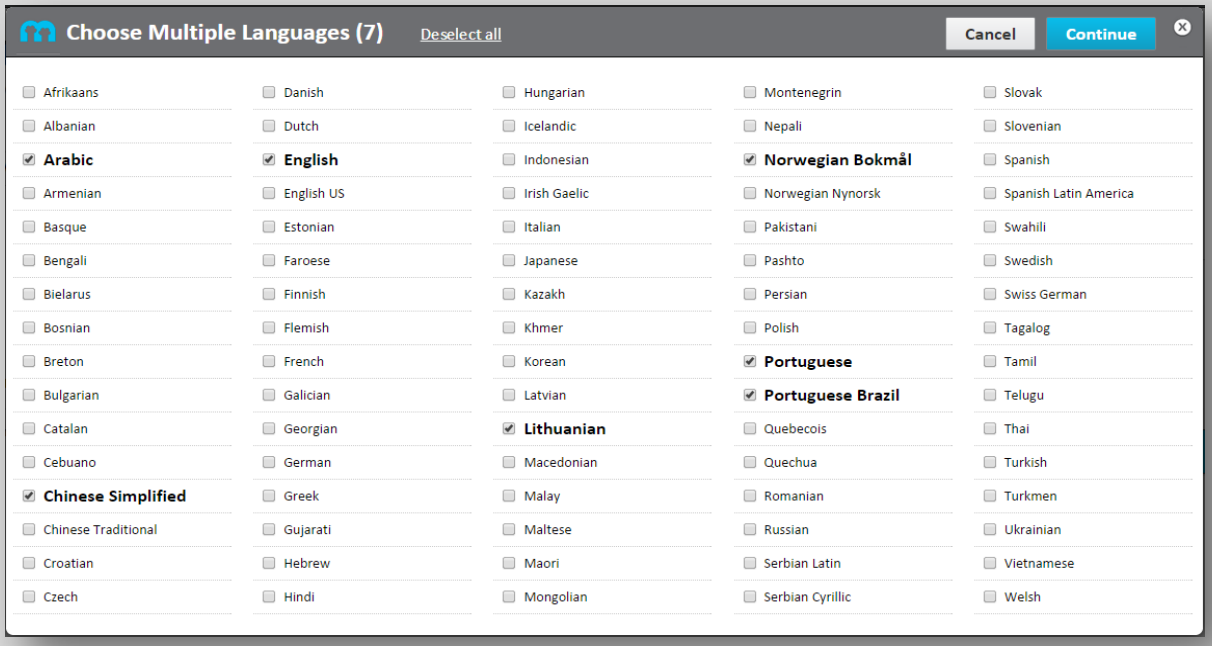
Always remember to choose the **language combination(s)** of your project before uploading the file(s) to translate, otherwise you will be asked to upload the file(s) again and the following message will be displayed:



If you need to select **more than one target language**, please use the Multiple languages? link.



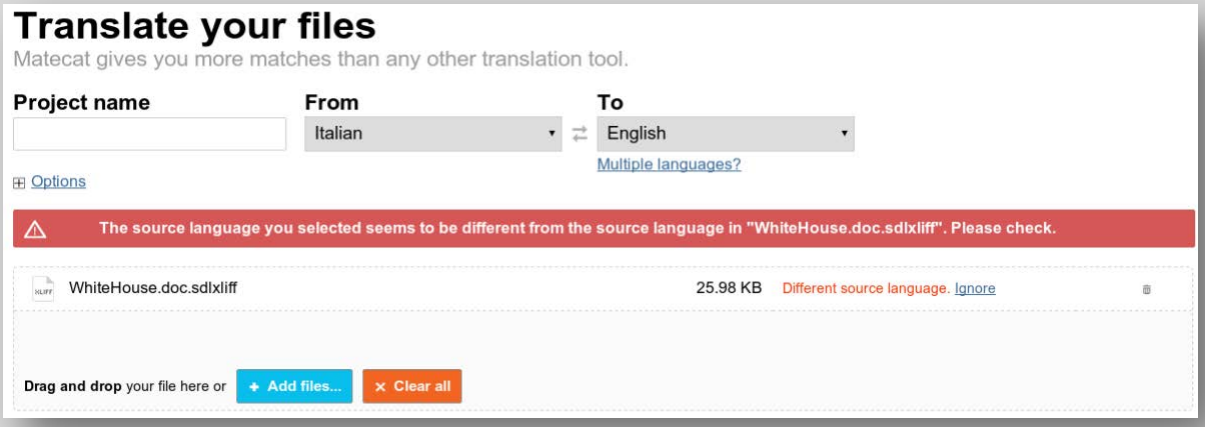
A box with the list of all supported languages will be displayed. Check the boxes of all languages you need to translate into, as shown in the screenshot below:



In order to upload the file(s) for translation you can:

- directly **drag and drop** them or
- **browse** for them by clicking on the **+ Add files…** button.

MateCat automatically **checks the source language** of your documents to make sure that it matches the one you selected. Otherwise, a warning message is displayed as in the screenshot below

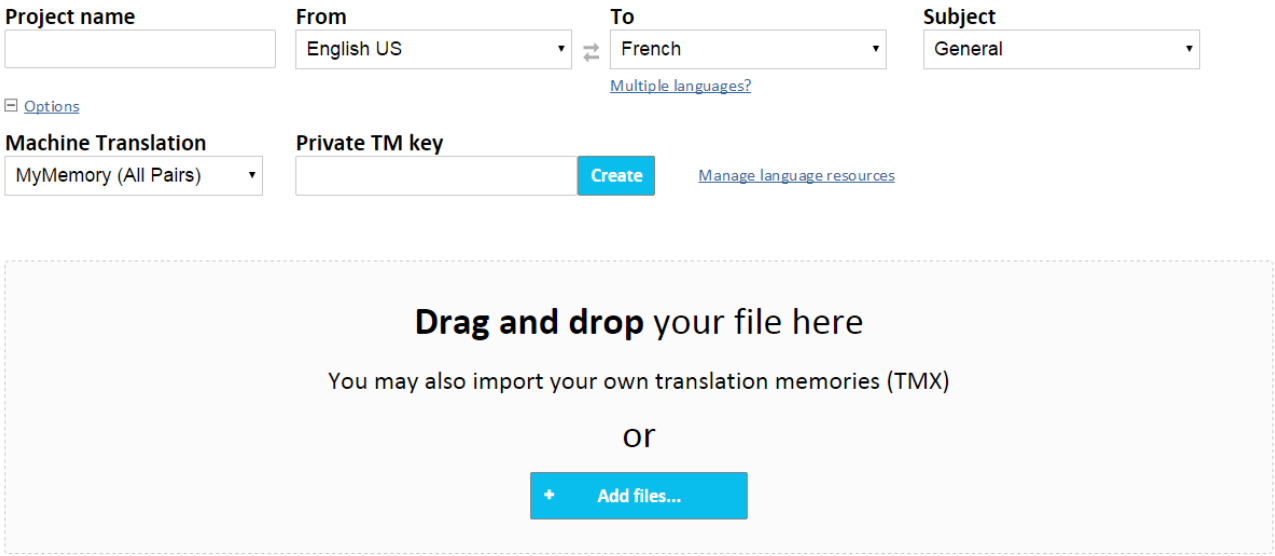


It also checks that the private TM key you associated with your project is indeed **valid**.

You can use the **dustbin** icon to delete a file from the list of files to translate or click on the **Clear all** button to delete all uploaded files from the list.

Translation Memory and Machine Translation

Clicking on the **Options** button on the MateCat home page opens a window, as in the screenshot below:



Here you can add and/or select **additional options** to be added to the project you are creating. For example, choose to get or not get machine translation suggestions from the MT server MyMemory, or add or create a private TM for your project.

For further details, refer to the Translation Toolbox section.

Machine Translation Servers

In MateCat, the best option is to select **MyMemory** which uses a combination of Google Translate and Microsoft Translator to provide machine translation suggestions. You can also disable machine translation, by selecting **NONE** from the dropdown menu.

Analysis Report

During creation of the project, when all files have been uploaded and all information has been correctly set on the MateCat home page, the **Analyze** button will be displayed. Click on the **Analyze** button to start the analysis.

The analysis page will open.

A progress bar shows you the progress on analyses of the files you uploaded for translation. Once the analysis is completed, the volume analysis page displays the **Analysis report.** It contains information about the number of words (**Payable, Total, New, Repeated**) for the entire project and for each file/job.

Payable words are marked in green and calculated by multiplying the words for each match type by the **payable percentage**:

*(1,903*0.3) + (112*0.6) + (144*0.3) + (9,493*0.85) = 8,750*

Assume you obtain the following analysis result:

	Payable	Total	New	Repetition	Internal Matches (75%-99%)	TM Partial (75%-99%)	TM 100%	TM 100% in context	Machine Translation
Payable Rate Breakdown			100%	30%	60%	60%	30%	0%	85%
12970 File details (1)	8,750	12,077	0	1,903	0	112	144	0	9,493

In detail, we have the indication of:

Payable words

Payable word count is the sum of the **weighted word count** for each match type multiplied by its **payable rate percentage**.

Total words

The **total word count** without leveraging any content from translation memory matches, repetitions or machine translation. This is similar to what Microsoft Word would give as a word count in a .doc or .docx file.

New words

All words found in segments that:

- **do not match** any fuzzy or complete match in the private TM Key and/or in the public TM;
- are **not repeated** in the project;
- **do not have a suggestion** from a machine translation.

Repetitions

Number of words of **identical** segments that occurs more than once throughout the project.

For example, imagine that we find the following segments in our translation:

1. My house is blue
2. My house is blue
3. My house is red

Segments 1 and 2 are **identical segments,** so they are counted as 4 repetitions.

Segment 3 is counted as an **internal match**.

Internal Matches

Internal matches are **similar segments** found in the document you are translating. For example, imagine that we find the following segments in our translation:

1. My house is blue
2. My house is red

MateCat recognises that segment 2 is similar to segment 1 (3 words out of 4 are identical) so the following would occur during translation:

1. You translate segment 1
2. The translation memory is **updated** with this translation
3. You open segment 2 for translation
4. A search is performed in the TM for segment 2 and a **75% fuzzy match** is found.

MateCat therefore counts four new words for segment 1 but, because a 75% fuzzy match has been found in the TM, the four words in segment 2 are counted as **Internal Matches** (in terms of weighted words, these are counted as 2.4, or 60% of 4).

Partial TM

In this case, similarity between the document to be translated and any **correspondences found in the translation memory** (fuzzy matches) are calculated.

For example, imagine that we have following segments in an EN>FR translation:

1. My house is blue
2. My house is red

and that the TM contains this:

Source	Target
My house is blue	Ma maison est bleue

For segment 1 (“My house is blue”), there will be a **100% match** in the translation memory. For segment 2 (“My house is red”), there will be a **75% fuzzy match** (segment 2 and the segment found in the translation memory only differ in terms of the colour of the house: blue/red –bleue/rouge, so 1 word out of 4).

100% TM

This is a 100% match between a segment in the source language found in the document to translate and an **identical sentence** found in the source language in the translation memory.

100% TM in context

This is **more than a 100% match**.

If you have a 100% context match (the corresponding label is “**100% CM**”), this means that both of the following **2 conditions** exist:

- the segment in the document has a **100% match** in the translation memory
- the segment in the document and the segment in the translation memory must both be **preceded by the same segment**

For example, if you have to translate the following segments in the same order:

1. My house is blue
2. My house is red

and you have a **100% match** in the translation memory for segment 2, you will actually receive a **100% CM** if segment 2 is also preceded by segment 1 in the translation memory.

This means that you can be certain that the 100% match is correct according to the **context** of the document you are translating.

Actually this does not mean that, in the list of segments in the translation memory, the segment in the translation memory is indeed preceded by the same segment as the one in the document. What it means is that **context information is stored in the translation memory as metadata**, so each segment stored in the translation memory contains the following information (not visible in the translation memory contents by the users):

- the source segment
- the corresponding translation
- the segment preceding the segment itself.

Click on the **Translate** button for each language pair to open the Translation Editor window and start translating. If you need to outsource it for translation to someone else, you should just send the **job URL** to your external resource. The URL is all you need to start translating.

Public Translation Memory and Translation Memory Key

When translating, your translations are saved as bilingual sentences in a so-called **translation memory (TM),** a sort of database containing bilingual segments (sentences, sets of words or single words already translated in one or more language pairs) which can be used during the translation of the same project or future projects.

So, when translating, if you find a segment that is identical or similar to another one that has already been translated, the translation memory suggests the wording of the already translated segment so that you can use it, partially or entirely, for translation of the new segment.

In MateCat, you can save and store your translations:

- in a **generic public translation memory** which can be accessed via API by all MateCat users.
- in a new or already existing **private TM assigned to one or more projects** or to a specific client that enables you to save your translation in a private translation memory that **cannot be accessed by external users.**

In MateCat, each private TM is identified by a unique alphanumeric value called **private TM Key.** It is the only value you need in order to associate your TM(s) with your project. It is like a password to access your private TM in the MateCat Translation Memory server.

If you save your translations in the private TM, they will be stored in your private translation memory only and you alone have **exclusive** use of and access to your contributions to that private TM.

If you save your translations in the public TM, they can be viewed as suggestions by all MateCat users. It would represent a **human contribution to a public TM** and would help **other MateCat users** during their translation jobs in MateCat.

TM backup and updates

Adding a private TM Key to your project means that the TM is **automatically updated during translation**, if previously set.

MateCat stores your TMs for you so that you do not need to **upload and download** the translated segments each time you translate. In this way, you have a **constant backup** of your translation memory and you can use it again in the future by just storing the private TM key and associating it with a new project, typing it in again if needed.[3]

How to add a Glossary

A glossary is a bilingual or multilingual database containing the translation of **terms**[4] in a special subject, field or area of usage. It is usually in Excel or CSV format and can help translators with the translations of **specific terms** that can be translated in several ways depending on the subject matter.

For example, imagine that we have to translate a file containing the following phrase from English to French:

My house is red.

In order to translate the file containing this phrase, we created a project for translating from English to French and associated a **private TM key** and an **EN>FR glossary**. In the glossary, we have the following term and its translation:

Source	Target
House	Maison

[3] It is not possible to download the translated segment for a specific project yet. We plan to develop such functionality in one of our next releases.

[4] Terms are words or set of words of a particular kind of language or branch of study.

Opening the phrase for translation, we have no suggestions from our translation memory but we have the translation of “house” in our glossary. In this case, MateCat suggests how to translate the term “house”.

In MateCat, **the glossary is included in the private TM key** associated with the project. If no private TM key is associated yet, a new one will be created once you insert the first glossary term.

Note: please refer to the section How to Use Glossary Terms in Your Translated Segment to know how to add a glossary term while translating.

CSV is the only supported glossary format. You just have to upload it to your private TM key.[5]

Why add a glossary for your project? If you upload a glossary, MateCat also looks up single terms. In the translation memory, this function is performed at **segment level**, whereas in the glossary it is performed at **term level**.

Translation Editor User Interface

Documents are translated in the **Translation Editor** window. This is the translator’s main interface. You can access this page:

- By clicking on the **Translate** button on the Analysis page
- directly from the corresponding previously generated/provided **job URL**

[5] The user cannot upload the glossary by him/herself yet. We plan to develop such functionality in one of our next releases.

The job URL contains information about the **job** you are translating. Below, you will find an explanation of the different elements of the URL.

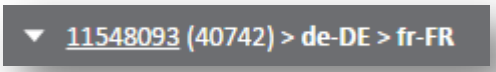
- www.matecat.com – **Base URL** for the MateCat proprietary version.
- /translate – This indicates that you are in the **translation editor page**.
- /Test – The **name** of the project you are translating.
- /en-US-it-IT – The **language pair** of your translation.
- /43505-0d5d8d53c774 – The **job ID** (43505) followed by the unique **job "password"** (in this case, 0d5d8d53c774).
- #21694626 – The **unique identifier of the segment** you are translating. This number changes every time you open a new segment.

The **Translation Editor User Interface** contains only one environment with all the information necessary for the translator to work on the job.

The interface contains the following bars and information:

o Top bar

Click on the logo to go to the **homepage** of the MateCat tool and create a new project.



The **name** of the job and **language pair** combination. Click on the arrow to see the list of the files and navigate through them.

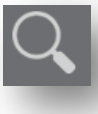
Click to **download the original** file(s).

Click to **download a preview of the translated file(s)**. The preview contains your translation and machine translation for the remaining untranslated segments.

The **Preview** button is replaced by the **Download translation** button once the translation is **complete** (**100% on the progress bar**).

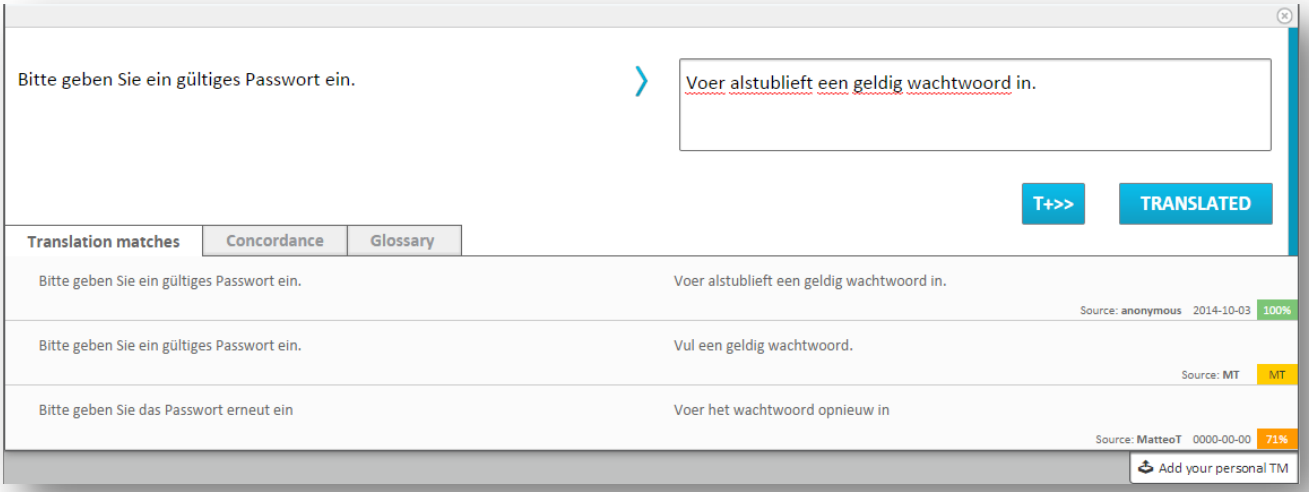
QA module – No errors reported.

QA module – **1 error message** reported. Click on the icon to go to the segment with the error and fix it.



Search tool – Perform a search in the **source or target text**. If necessary, you can replace the term you searched for with another one.

o Editor window

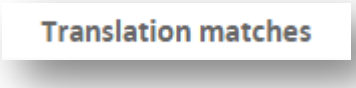

Bitte geben Sie ein gültiges Passwort ein.
Voer alstublieft een geldig wachtwoord in.
T+>>
TRANSLATED
Translation matches
Concordance
Glossary
Bitte geben Sie ein gültiges Passwort ein.
Voer alstublieft een geldig wachtwoord in.
Source: anonymous 2014-10-03 100%
Bitte geben Sie ein gültiges Passwort ein.
Vul een geldig wachtwoord.
Source: MT MT
Bitte geben Sie das Passwort erneut ein
Voer het wachtwoord opnieuw in
Source: MatteoT 0000-00-00 71%
Add your personal TM

Click to **close the segment** in draft status, without saving it in your TM

Copy source to target button. Click or use the shortcut ALT+CTRL+i to copy the source text to the target box

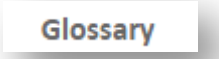
Translated and go to next untranslated button. Click or use the shortcut CTRL+SHIFT+Enter to **confirm your translation of the current segment and to move to the next untranslated segment**.

Translated button. Click or use the shortcut CTRL+Enter to **confirm your translation and save it in the translation memory**

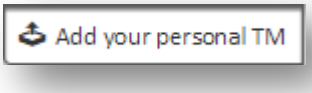


Translation matches tab. This tab shows the user the **TM matches and/or MT suggestions** for the current segment.

Concordance tab. Open this tab or use the shortcut ALT+c to **search the translation memory,** both public and private, if any, for a particular word, word sequence or phrase



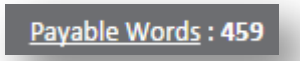
Glossary tab. Open this tab to **check suggestions from a previous glossary** included in the private TM key which is associated with the project or to **add new terms to the glossary**.



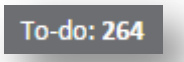
Click to **add your personal TM** to the ongoing translation.

o Bottom bar

The **progress bar** indicates the **real time status** of your translation.

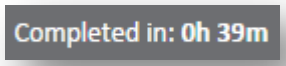


Number of **payable (weighted) words** for the entire job. Clicking on **Payable Words** button will take you to the Volume Analysis page of the project.

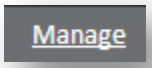


Number of words left to translate for the entire job.

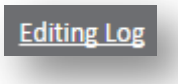
The **estimated speed of your translating**. Based on the average time spent on the last 10 translated segments, MateCat calculates **the hourly turnaround** for the job if you are currently working on the translation. After one hour of inactivity, this information is no longer available.



The **estimated time left to complete** the translation. Based on the time spent on the last 10 translated segments, MateCat calculates **how much time is needed to complete the translation** if you are currently working on the translation. After one hour of inactivity, this information is no longer available.

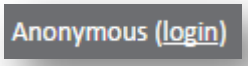


Link to access your own **Management Panel**



Link to access the **Editing Log page**, a **statistical page** about the current project. (e.g. percentage of MT matches, percentage of TM matches, average post-editing efforts, etc.)

Link to access the manual



Information about the **account you are logged in with**. Click on login to sign in using Google.

Source text and target text are displayed **side by side**.

Click on any segment in the target column to open it for translation.

What if I need to stop translating and switch off my computer? In MateCat your translation is **automatically saved** in the document when you edit the segment, and in the translation memory, if any, when you click on the **Translated** or **T+>>** buttons. If you need to stop translating and switch off your computer, all you need to do is **open the** job URL **again** in your browser. MateCat automatically opens the last segment you edited. If you created the project yourself, by logging in with your **Google account,** you can recover the job URL from your **Management Panel**. If you were not logged in during the creation of the project or if you did not create the project by yourself, remember to store the job URL in a safe place in order to recover it again easily.

Translation Toolbox

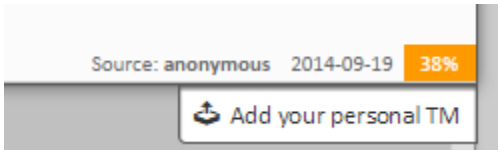
Manage your language resources

You can store all your private TMs in MateCat and assign them to your projects when creating them or during translation.

Adding a TM to a project

When creating a project, click on “Options” and then on “Manage language resources”; a new window slides in from the right of your screen. Click on “+ Add a TM” to open up the box to add your TM (see below).

You can do the same while translating: click on “Add your personal TM” and a new window slides in from the right of your screen. Click on “+ Add a TM” to add your private TM.



If you already have a private key (that is, an alphanumeric string that identifies your private TM), you can paste it in the first input box, otherwise click on “Create”. Optionally, you can enter a short description to help you remember what the TM is for (e.g. “ACME bird seeds”). When you are done, click on “Confirm” and your private TM is created.

Importing your TMX files

After you create a private TM as described above, you can import your TMX[6] files into it to re-use your previously translated content.

To import a TMX to your private TM, click on “Import TMX”, select the TMX file to import and confirm. Note that a private TM on MateCat can contain any number of languages, so you can put all languages in a single private TM for a client, product, subject or whatever is more convenient for you.

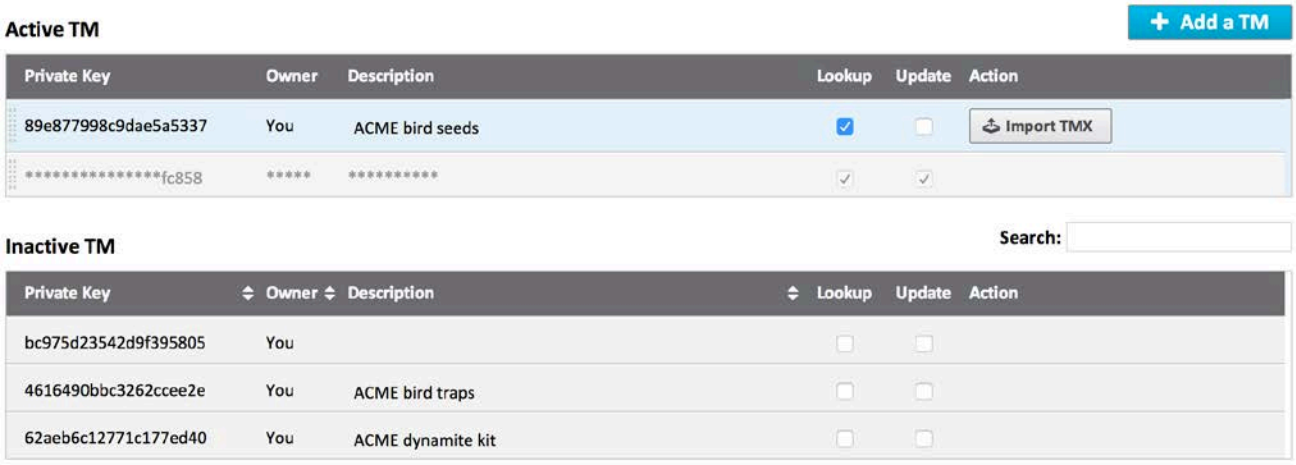
Active and Inactive TMs

In your Language resources panel you will see two separate lists:

- Active TMs are the translation memories that you have assigned to the project you are creating or to the project you are working on;
- Inactive TMs are the translation memories that you have saved in your profile on MateCat and that you can add to a project by selecting the “Lookup” checkbox and/or “Update”.

You can activate a TM by selecting the Lookup or Update checkbox.

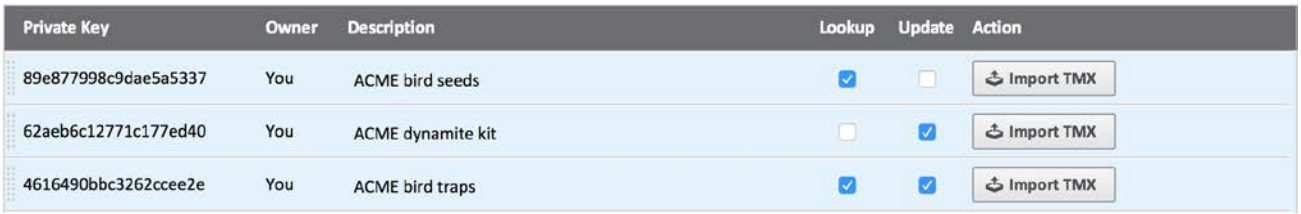
[6] TMX is one of the most common translation memory formats. It is the main translation exchange format.



Look up and Update

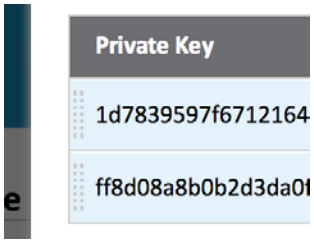
For any TM, you can decide whether you want to use it to look up translation matches or to also save your translations to it. To do that, select one or both of the "Lookup" and "Update" checkboxes.

- Only "Lookup" selected: the translation memory is only used to look up translation matches from your private TM;
- Only "Update" selected: the translation memory is only used to save your translated segments in the private TM;
- Both "Lookup" and "Update" selected: the translation memory is used both to look up translation matches and to save your translated segments in the private TM.



Adding multiple private TMs to your project

You can add any number of private TMs to your project. When you add more than one private TM to your project, you can order them by dragging and dropping each TM to the desired position.



The translation matches that MateCat presents for each segment are ranked based on their fuzzy match score. However, when two matches coming from two different TMs have the same fuzzy match score, they are ranked based on the priority of the private TMs in your "Active TM" list.

Private TMs from other users

When you are working on a project created by another user (e.g. your PM) or you are working with a colleague, you may see translation memories that do not belong to you: they are greyed out and covered with asterisks.

These are private TMs that another user added to the project and chose to share with you. Only the owner of those TMs can decide whether to add them to a project or to make them available to you in "Lookup" and/or "Update" mode.

Sharing your private TMs

You can share your private TMs with other users simply by sending them the private key (e.g. 89e877998c9dae5a5337). All they will need to do is to copy and paste the private TM into the private key input box that is opened when clicking on "+ Add a TM".

Note: When you share a private key, you are in fact giving full access to that translation memory. Users with your private key can use the private TM for their projects, update it and even delete it.

Working Offline

MateCat is designed for users that have access to the internet. However, if you need to work offline on a project that you started with MateCat, you can export it from MateCat and continue with a desktop CAT tool.

Access the Manage panel, and click on Export SDLXLIFF, open the file with a CAT tool that supports this format (SDL Trados, MemoQ, OmegaT and many more).

You can also use the Export feature to perform external QA checks with software such as ApSIC Xbench or Verifika.

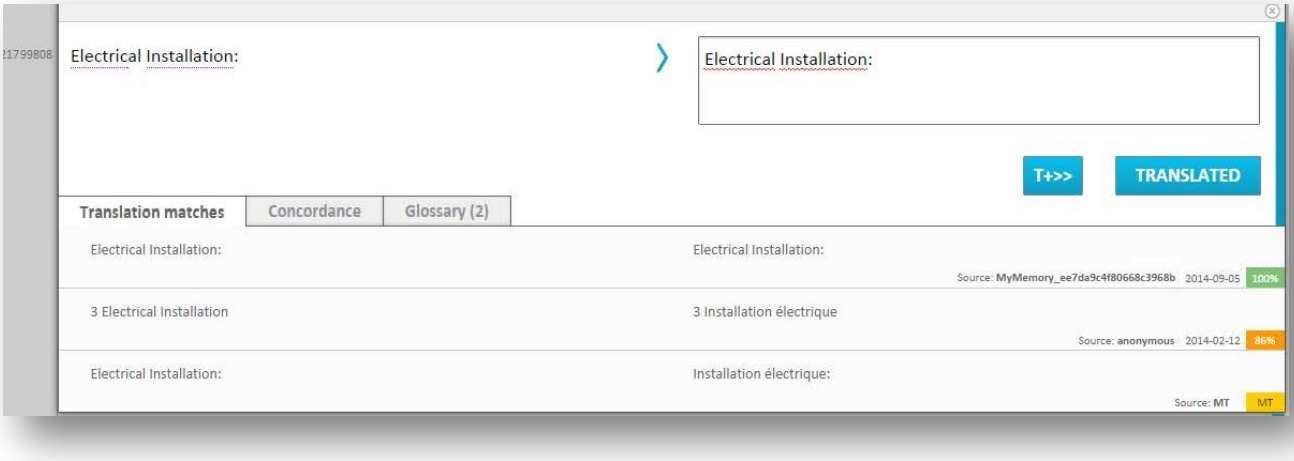
Using TM matches, MT suggestions and Glossary terms

When you open a segment for translation it will automatically populate with the first TM/MT match available for that segment.

You will receive **TM suggestions** if you add a private TM key to the project or in the case of fuzzy matches found in the public translation memory. They will be displayed as first suggestions just below the segment, under the **Translation matches** tab.

> If you have a personal translation memory, you can **add it to the project**. Please refer to the *Manage your language resources* section**.**

On the right, the tool gives information about the **source of the match**, the **creation date** and the **match percentage**, as shown in the screenshot below:



In this example, we have:

- a **100% match** in the private TM key that we added to the project (the private TM key is shown as the **source** of the suggestion) for the first suggestion; the segment is therefore automatically populated with this suggestion.

> When you add a private TM to your project, MateCat automatically associates a **name** with the private TM key in order to recognise the source of the suggestion while translating. It has the MyMemory_XXXX format, where XXXX is a **unique alphanumeric value** for each private TM key.

- an **86% match** from the public TM (in this case the **source** is "anonymous") for the second suggestion
- the **machine Translation suggestion** (the label of the source of the suggestion is **MT**) for the third suggestion.

Depending on the source of the match and on the match percentage, the following labels will be displayed:

If you have a **100% match** from the TM.

If it is a **MT** suggestion.

The corresponding match percentage in case of a **fuzzy match** from the TM.

If you are taking part in a collaborative project, you can also receive **real time suggestions** from the other translators working on the same project. This means that if more than one translator is working on the same project because it has been split among **several translators**, these translators share a private TM key and each of them receives suggestions for segments already translated within the same project. In this case, the source of the suggestion will have the name of the **shared** private TM key.

You will see the suggestions ranked from the highest percentage of matching to the lowest. The MT suggestions correspond to an **85% match by default**. If there are no higher percentages of matching, you will see the MT as the first suggestion.

To **select and enter one of the suggestions** into the translation input field you can:

- Use the **shortcut** ALT+CTRL+[n] for Windows or ALT+CMD+[n] for Mac OS X where [n] could be 1, 2 or 3, depending on whether you would like to add respectively the first, second or third suggestion.
- **double-click on the match** you would like to select

If you hover over a TM match, a **dustbin icon** appears on the right. Click on it to **remove the TM match from the translation memory**.

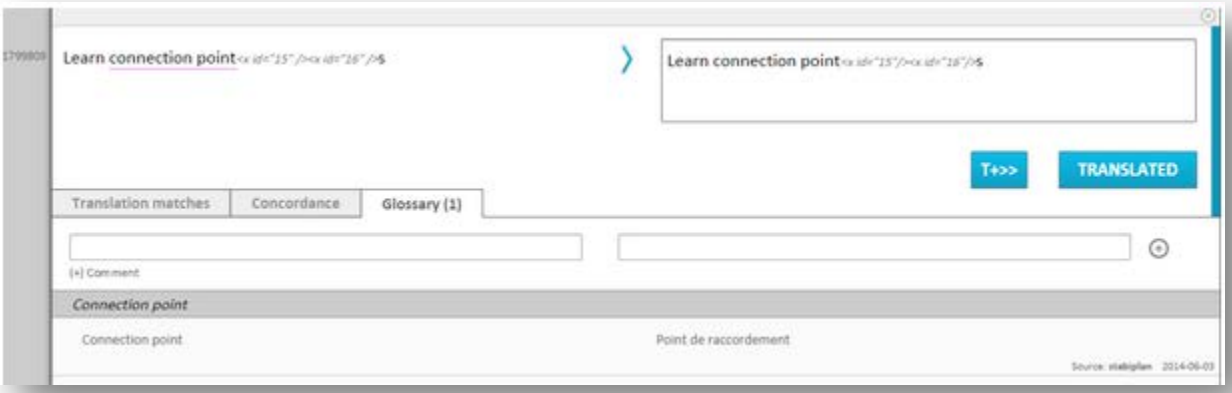
Note: click on the dustbin icon only if you would like to **delete the match from the translation memory** because the target text is not the correct translation of the source text. Do not delete it if you think it is not the correct suggestion for the segment you are translating.

If you would like to **copy the source segment to the target input area**, click on the arrow between the source and target segment ❭ or press ALT+CTRL+I

How to Use Glossary Terms in Your Translated Segment

The **Glossary** tab provides **suggestions from the glossary** previously included in the private TM key which is associated with the project.

Every time you encounter this term in the text, it is displayed by the glossary and the **term in the segments underlined**, as shown in the screenshot below:



The translations of these terms are listed under the **Glossary tab** which indicates in brackets the number of terms in the segment found in the glossary.

Each glossary suggestion also gives information about the **source of the translation** and the **creation date**.

If you hover over a suggested term, a **dustbin icon** appears on the right. Click on it to **remove the term from the glossary**.

You can also **add new terms** to the glossary:

- Add source and target terms in the corresponding windows
- Add comments, if needed, by clicking on the **(+) Comment** button

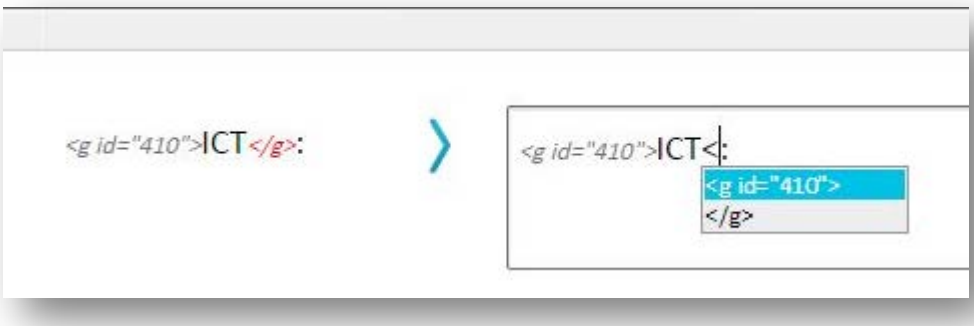
- Click on the ⊕ button to add the term to the glossary

Note: If no glossary is added to the project, you **create your own new glossary** by creating a new term, and, if no private TM key had been previously associated with the project, a **new private TM key will be generated automatically** and added to the project.

Working with Text and Tags

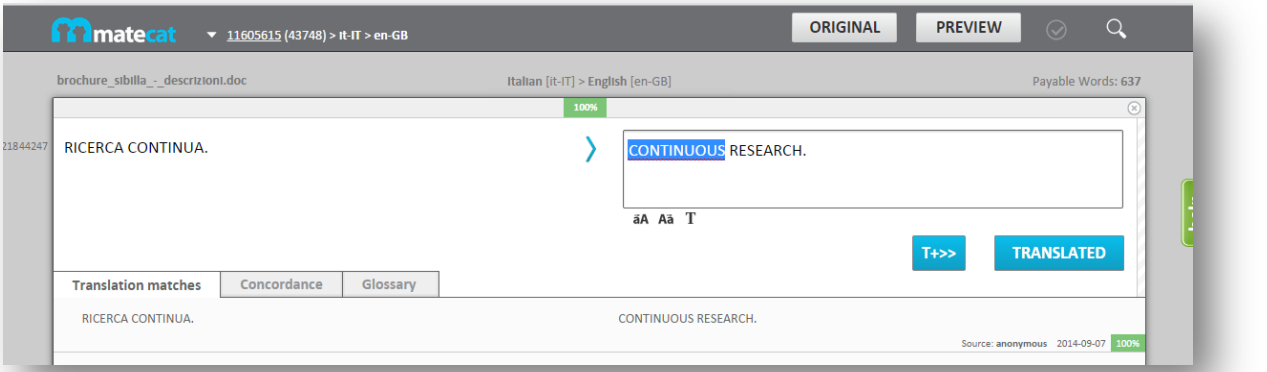
o Tag autocompletion

When you need to input a tag into your translation, enter the symbol < and a list of the tags in the source text is displayed. Select the one you need and click Enter to add it to your translation.



o Formatting features

MateCat uses a simple functionality to **change the case of words** in the editing area. Just select one or more words and **three buttons** will appear to allow you to select the formatting you want and switch the case (**upper case**, **lower case** and **capitalise first letter**).



o Adding line breaks and hidden text

If you need to add **line breaks** in the target segment:

- place your cursor where you want the text to break to a new line,
- press **Shift+Enter** and the segment will divide into two separate lines.

The same number of lines will also be retained in the target document.

MateCat uses **symbols for hidden characters** such as **tabs**, **non-breaking spaces** and **line breaks**, which are respectively →| , · and ↵. If one of them is present in the source text, you must reproduce it in the target segment using the corresponding key or shortcut (the Tab key on your keyboard to add a tab, the Shift+Enter shortcut to add a line break and Ctrl+Shift+Space shortcut to add a non-breaking space). Do not copy and paste them from the source text.

o Placeables and untranslatable text

Currently, we cannot manage placeables and untranslatable parts. We plan to develop such functionality in one of our next releases.

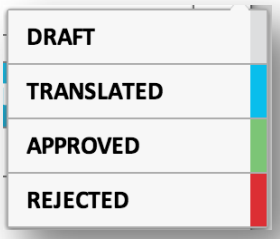
Confirming a Segment

Once your translation has been entered or the pre-translated segment edited, click on the **Translated** button to **confirm your translation and move to the next segment**. Alternatively, you can use the keyboard shortcut **Ctrl+Enter**.

If you want to **confirm the translation of the current segment and go to the next untranslated segment**, click on the **T+>>** button or use the keyboard shortcut **CTRL+SHIFT+Enter**.

The segment status bar colour will change from white to blue. In this way you **save the segment and update the translation memory**. If you do not click on the **Translated** button, and move manually on to the next (or press CTRL+Down) or to the previous (or press CTRL+Up) segment, your translation will be saved in the file but **the translation memory will not be updated**.

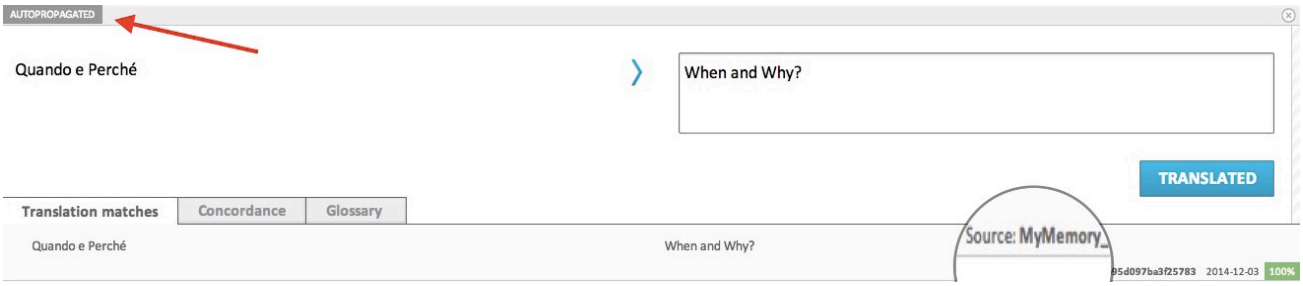
The bar to the right of each segment indicates its **status** using colours. Click on it to display the following options and to set/change the **status of a segment**:



Autopropagation

When you approve a segment during translation, MateCat **automatically populates** your translation with **all segments having the same source** within the same project.

They will be labelled as **AUTOPROPAGATED** as shown in the screenshot below



Your translation **will appear in all segments having the same source** but the status of the segments will remain "**untranslated**". You must approve them manually by clicking on the **Translated** button (or use the CTRL+Enter shortcut).

Concordance Search

Concordance searching allows you to search the translation memory, both public and private, if any, for a particular word, word sequence or phrase.

The search will find and display all previously translated segments in the private TM Key associated with the project and in the public TM containing that word, those words or that phrase.

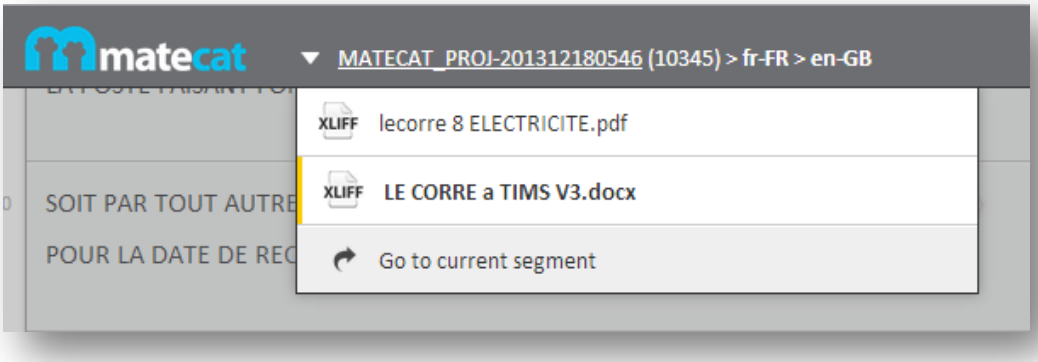
You can perform a concordance search by entering or pasting the text to check in the Concordance tab on the Translation Editor page. Alternatively, you can select the words you would like to check and use the following shortcuts which will automatically open up the Concordance tab:

- on Mac OS X: ALT+c
- on Windows: ALT+c

Navigating through Files

If you have uploaded **more than one file** for translation **in the same project**, you can **click on the job name at the top of the page to display a list of the files**. Clicking on the file name, you will be redirected to the beginning of that file.

Click on **Go to current segment** to display the segment you are working on.


matecat
MATECAT_PROJ-201312180546 (10345) > fr-FR > en-GB
XLIFF lecorre 8 ELECTRICITE.pdf
XLIFF LE CORRE a TIMS V3.docx
Go to current segment
SOIT PAR TOUT AUTRE
POUR LA DATE DE REC

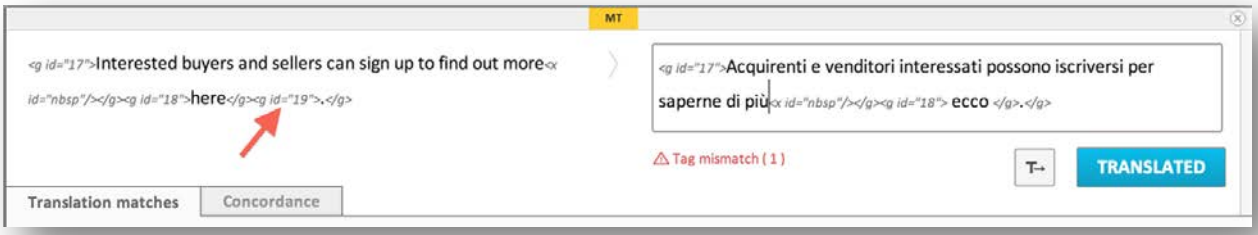
QA Messages

At the top right of the MateCat window, you will see an icon which warns of any potential QA errors in the translation. The warning sign contains a red circle indicating how many errors have been found in the project. None of these errors prevents you from downloading the translated file.

	The project contains one QA error.

Click on the icon to open the segment with the error. A message indicates the nature of the error: tag mismatch and/or translation conflict.

o Tag Mismatch

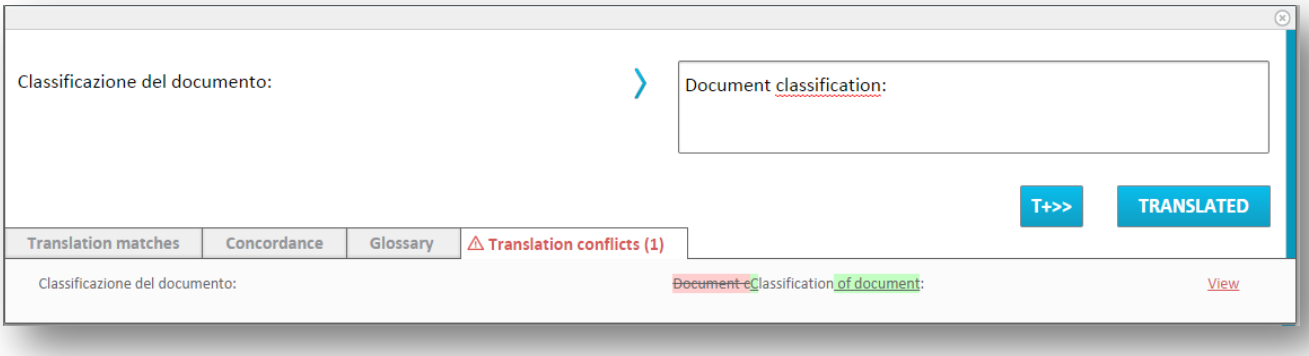


The tag <g id="19"> in the source segment, indicated by the red arrow, **was not added to the target segment**. This generated a tag mismatch. In order to fix the error, you can:

- copy and paste the tag from the source segment to the target and place it in the same position as in the source
- click on the tag in the source segment and drag and drop it to the target segment
- enter the symbol < to display the list of the tags in the source text. Select the one you need and click **Enter** to add it to your translation.

IMPORTANT NOTE: If you don't fix the tag mismatch error before clicking on the Preview or Download button in the top bar of the translation editor window, the strings "UNTRANSLATED_CONTENT" will appear in the segments where the error is present in the downloaded file(s).

o Translation conflicts

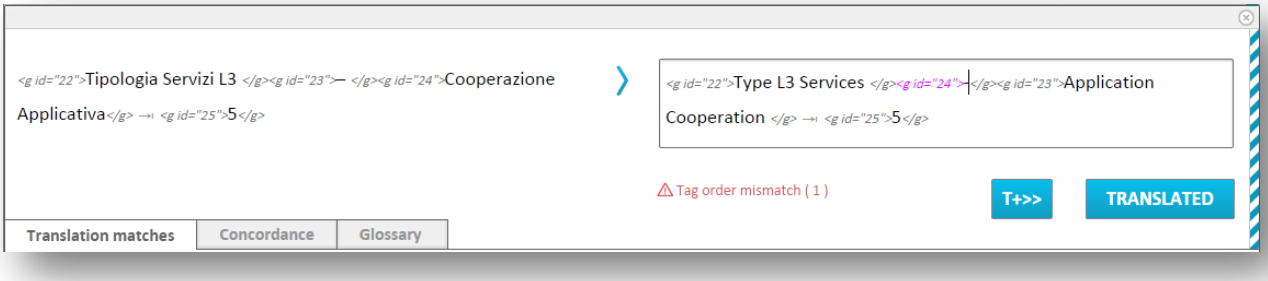


MateCat warns the user when **two or more different translations are inserted for the same source segment in the same project**. MateCat uses **track changes** to highlight the differences. To fix this error, choose one of the suggested matches or click on the **View** button to check and correct the translations.

……

The following two types of error messages will not appear in the warning sign. They will be displayed at the segment level only.

o Tag Order Mismatch



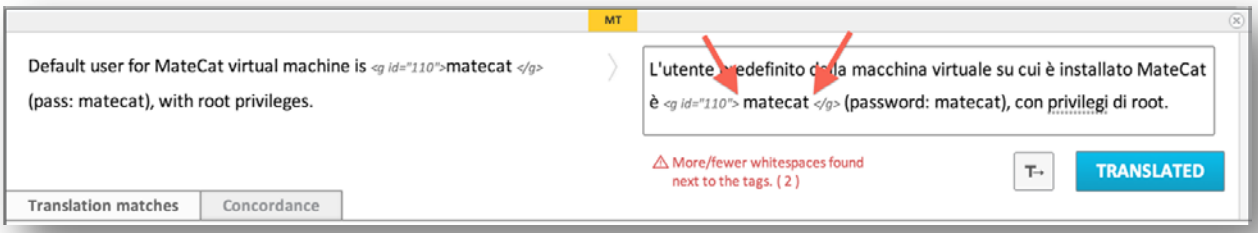
MateCat also displays a **warning** when tags are not in the **same position** as in the original segment. Tags in the wrong order will be coloured in pink.

In the example above, tag <g id=”24”> in the target segment is not in the same position as in the original. In order to fix the error, you can:

- click on the tag coloured in pink in the target segment and drag and drop it to the correct position in the phrase
- enter the symbol < in the tag’s correct position to display the list of the tags in the source text. Select the one you need and click **Enter** to add it to your translation. Then delete the tag in the wrong position.
- copy and paste the tag from the source segment to the target and place it in the correct position. Then delete the tag in the wrong position.

IMPORTANT NOTE: If you don’t fix the tag order mismatch error before clicking on the Preview or Download button in the top bar of the translation editor window, in most cases the translated document will still be created normally, but it’s possible that the original text will be displayed in the downloaded file instead of the translation.

o Whitespaces mismatch

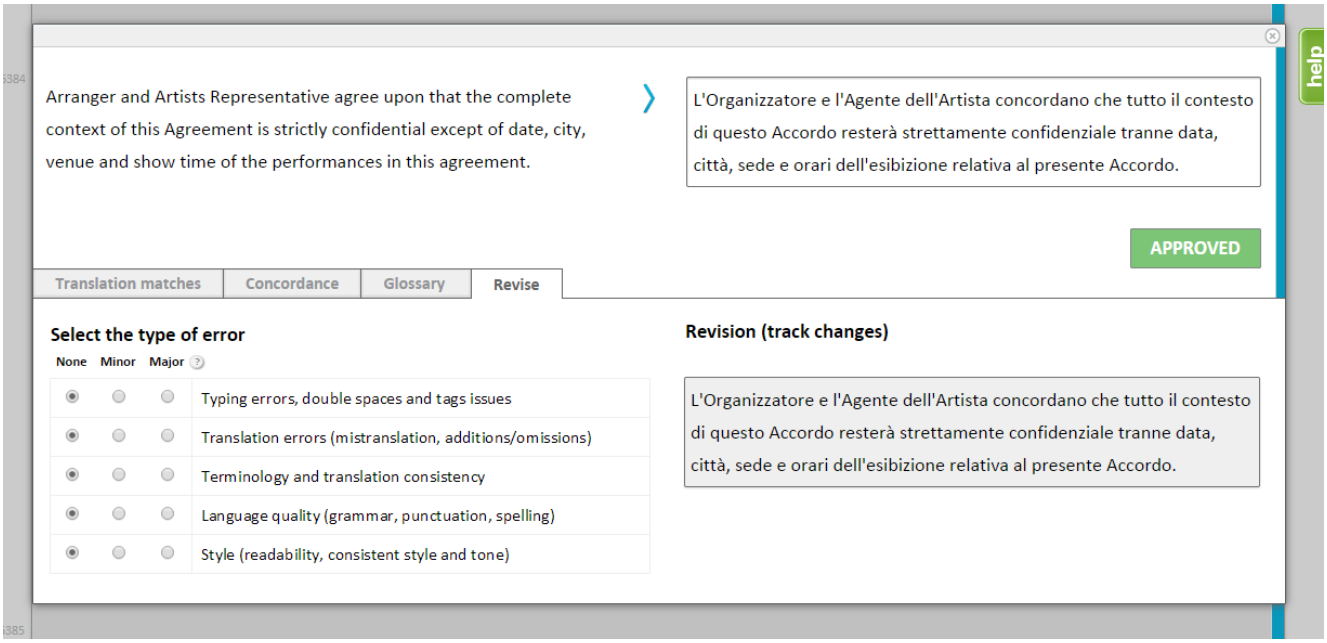

MT
Default user for MateCat virtual machine is <g id="110">matecat </g>
(pass: matecat), with root privileges.
L'utente predefinito della macchina virtuale su cui è installato MateCat
è <g id="110"> matecat </g> (password: matecat), con privilegi di root.
More/fewer whitespaces found
next to the tags. (2)
TRANSLATED
Translation matches
Concordance

MateCat warns the user when **more or fewer spaces** are found in the target segment next to tags. In the image above, there are two extra whitespaces indicated by the red arrows. To fix the error, delete the extra spaces between the tags and the word.

Revision in MateCat

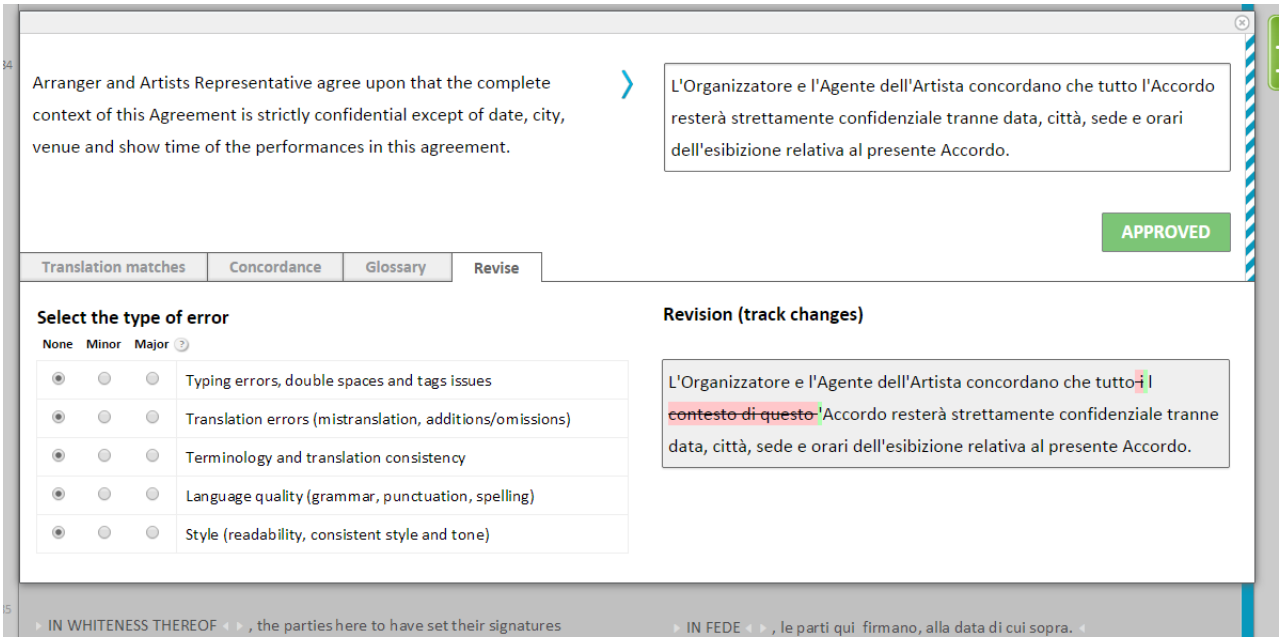
You can revise a project in MateCat during the translation phase or after the translation is completed.

Click on the Revise link at the right bottom corner of the Translate page to enter the revision mode.


Arranger and Artists Representative agree upon that the complete context of this Agreement is strictly confidential except of date, city, venue and show time of the performances in this agreement.
L'Organizzatore e l'Agente dell'Artista concordano che tutto il contesto di questo Accordo resterà strettamente confidenziale tranne data, città, sede e orari dell'esibizione relativa al presente Accordo.
APPROVED
Translation matches
Concordance
Glossary
Revise
Select the type of error
None
Minor
Major
Typing errors, double spaces and tags issues
Translation errors (mistranslation, additions/omissions)
Terminology and translation consistency
Language quality (grammar, punctuation, spelling)
Style (readability, consistent style and tone)
Revision (track changes)
L'Organizzatore e l'Agente dell'Artista concordano che tutto il contesto di questo Accordo resterà strettamente confidenziale tranne data, città, sede e orari dell'esibizione relativa al presente Accordo.
help

o Revision process

- If the existing translation does not require revision, just click on ***Approved*** (or press CTRL+Enter) and you will go to the next Translated segment to be revised.
- If the existing translation needs to be edited, edit the translation in the target window. You'll see the changes you make in real time in the Revision (track changes) window. Once you finish editing the segment, select the type of issues you found in the translation, then click on the Approved button (or press CTRL+Enter) to go to the next Translated segment to be revised.

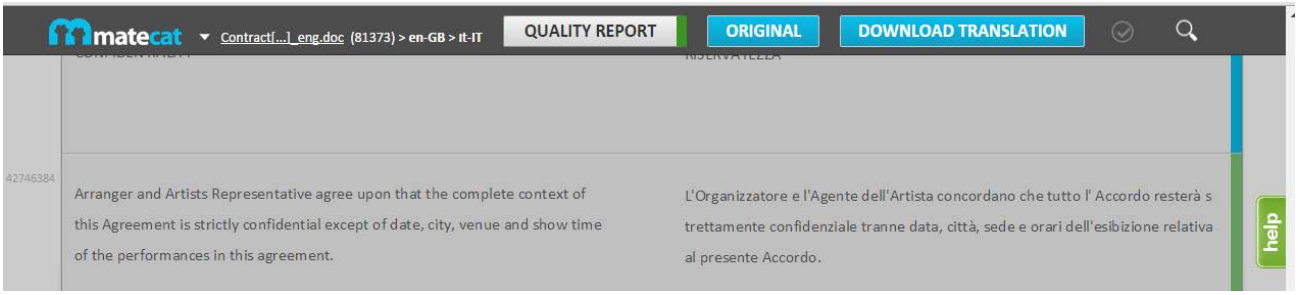

Arranger and Artists Representative agree upon that the complete context of this Agreement is strictly confidential except of date, city, venue and show time of the performances in this agreement.
L'Organizzatore e l'Agente dell'Artista concordano che tutto l'Accordo resterà strettamente confidenziale tranne data, città, sede e orari dell'esibizione relativa al presente Accordo.
APPROVED
Translation matches
Concordance
Glossary
Revise
Select the type of error
None
Minor
Major
Typing errors, double spaces and tags issues
Translation errors (mistranslation, additions/omissions)
Terminology and translation consistency
Language quality (grammar, punctuation, spelling)
Style (readability, consistent style and tone)
Revision (track changes)
L'Organizzatore e l'Agente dell'Artista concordano che tutto~~-i~~ l ~~contesto di questo~~ 'Accordo resterà strettamente confidenziale tranne data, città, sede e orari dell'esibizione relativa al presente Accordo.
IN WHITENESS THEREOF , the parties here to have set their signatures
IN FEDE , le parti qui firmano, alla data di cui sopra.

You can also move from one segment to the other by clicking on the segment you want to revise. However, if you edit a segment and you do not click on ***Approved***, your translation will not be saved. In order to complete and save your revision you will have to go back to all segments that are not marked in green, open the editing box and confirm the translation by clicking on the ***Approved*** button (or by pressing CTRL+Enter).

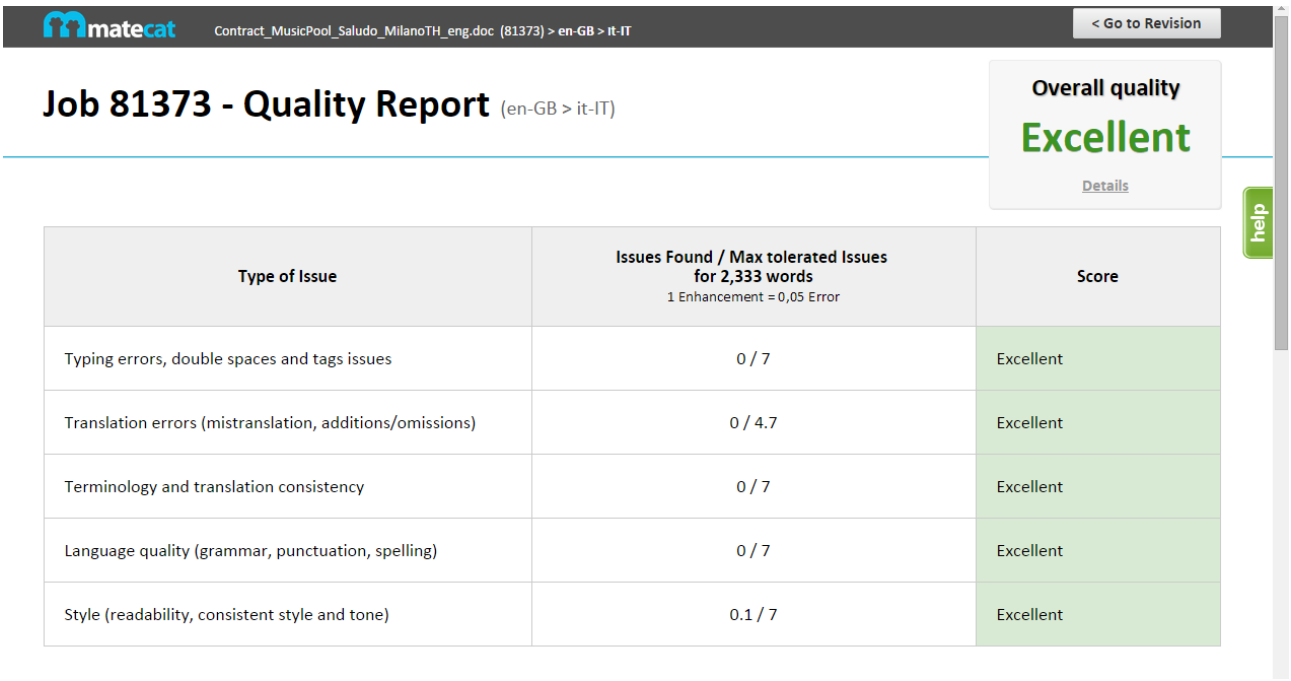
o Quality Report

While you revise the segments, the sidebar of the **QUALITY REPORT** button on the Revise page will change color in real time according to the ongoing result of the revision based on these values:

- Green= Excellent, Very Good, Good
- Yellow= Acceptable
- Orange= Poor
- Red= Fail

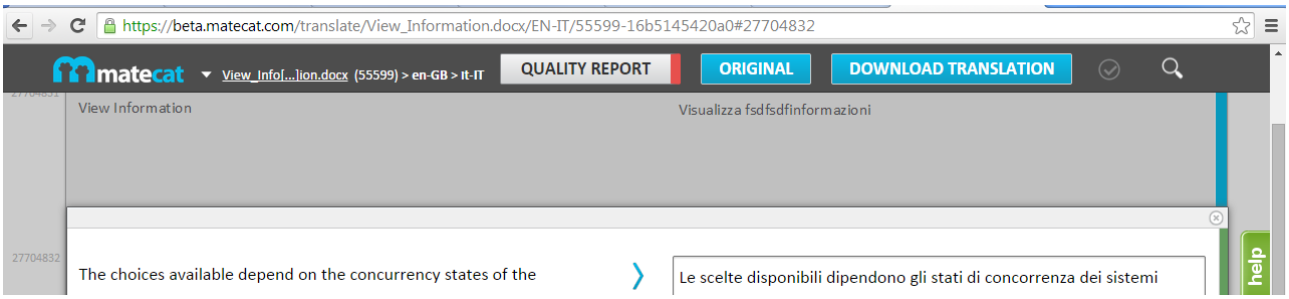

matecat
Contract[...]_eng.doc (81373) > en-GB > it-IT
QUALITY REPORT
ORIGINAL
DOWNLOAD TRANSLATION
42746384
Arranger and Artists Representative agree upon that the complete context of this Agreement is strictly confidential except of date, city, venue and show time of the performances in this agreement.
L'Organizzatore e l'Agente dell'Artista concordano che tutto l' Accordo resterà s trettamente confidenziale tranne data, città, sede e orari dell'esibizione relativa al presente Accordo.
help

Clicking on this button, you'll see the total quality report for the job revised and the total issues found per category.



Type of Issue	Issues Found / Max tolerated Issues for 2,333 words 1 Enhancement = 0,05 Error	Score
Typing errors, double spaces and tags issues	0 / 7	Excellent
Translation errors (mistranslation, additions/omissions)	0 / 4.7	Excellent
Terminology and translation consistency	0 / 7	Excellent
Language quality (grammar, punctuation, spelling)	0 / 7	Excellent
Style (readability, consistent style and tone)	0.1 / 7	Excellent

This button is also shown on the Translate page **only** if the Overall Quality is Acceptable, Poor or Fail.



o Finalizing

The revision is complete once all segments have been revised and marked as approved using the ***Approved*** button (or by pressing CTRL+Enter). At this point, the progress bar on the bottom of the screen will be all green.

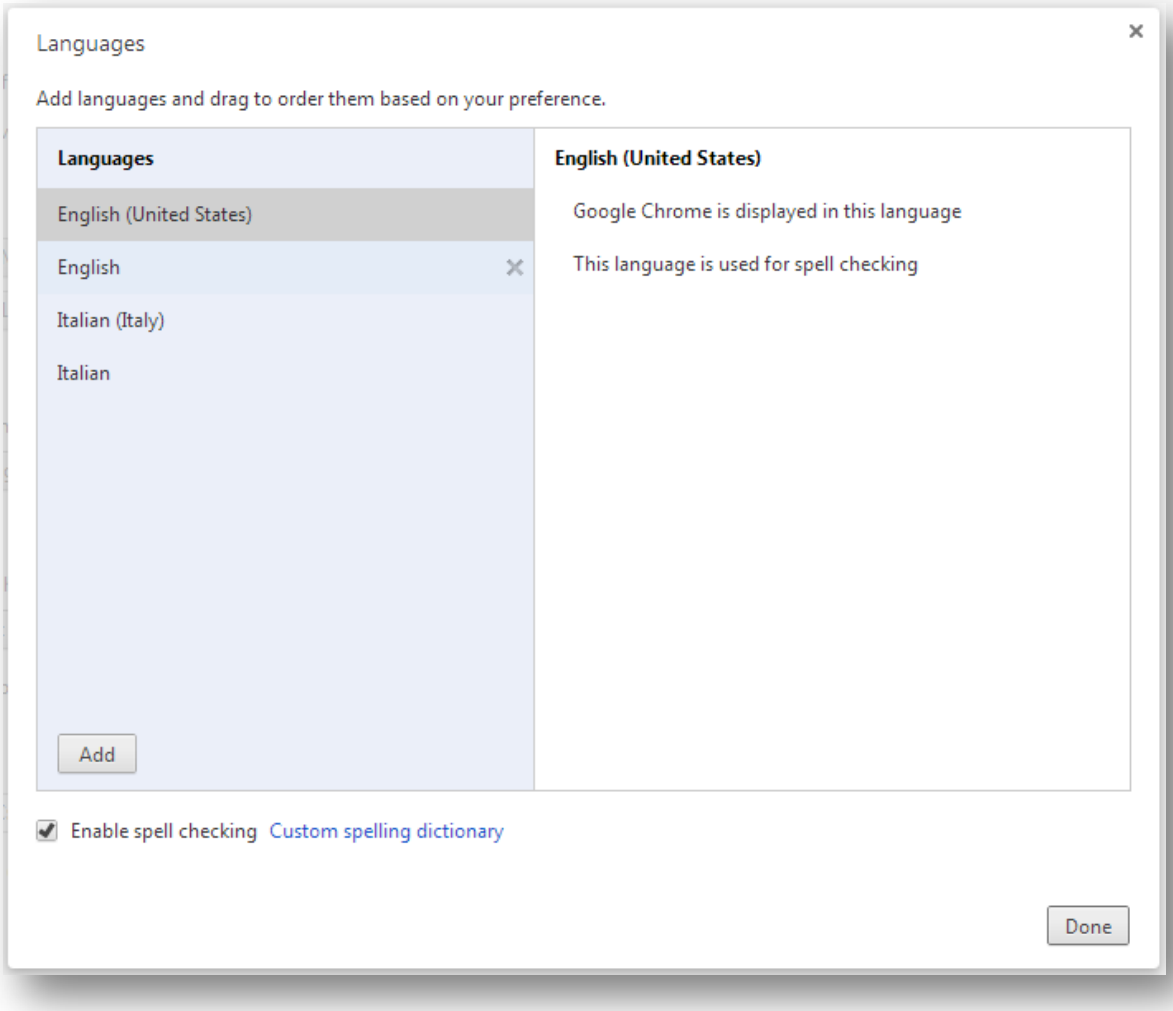
Make sure to proofread the file you download before delivering it to your client because some formatting cannot be corrected directly on MateCat.

Spell-checking

MateCat uses your browser's spell-checker.

To enable the **Google Chrome spell-checker** for your target language, follow these steps:

1. Enter the following string in the Google Chrome address bar: **chrome://settings/languages**
2. Click on Add and add your **target language**;
3. Select **Enable spell checking** in the dialogue window that opens.



Spell-checking enabled for US English

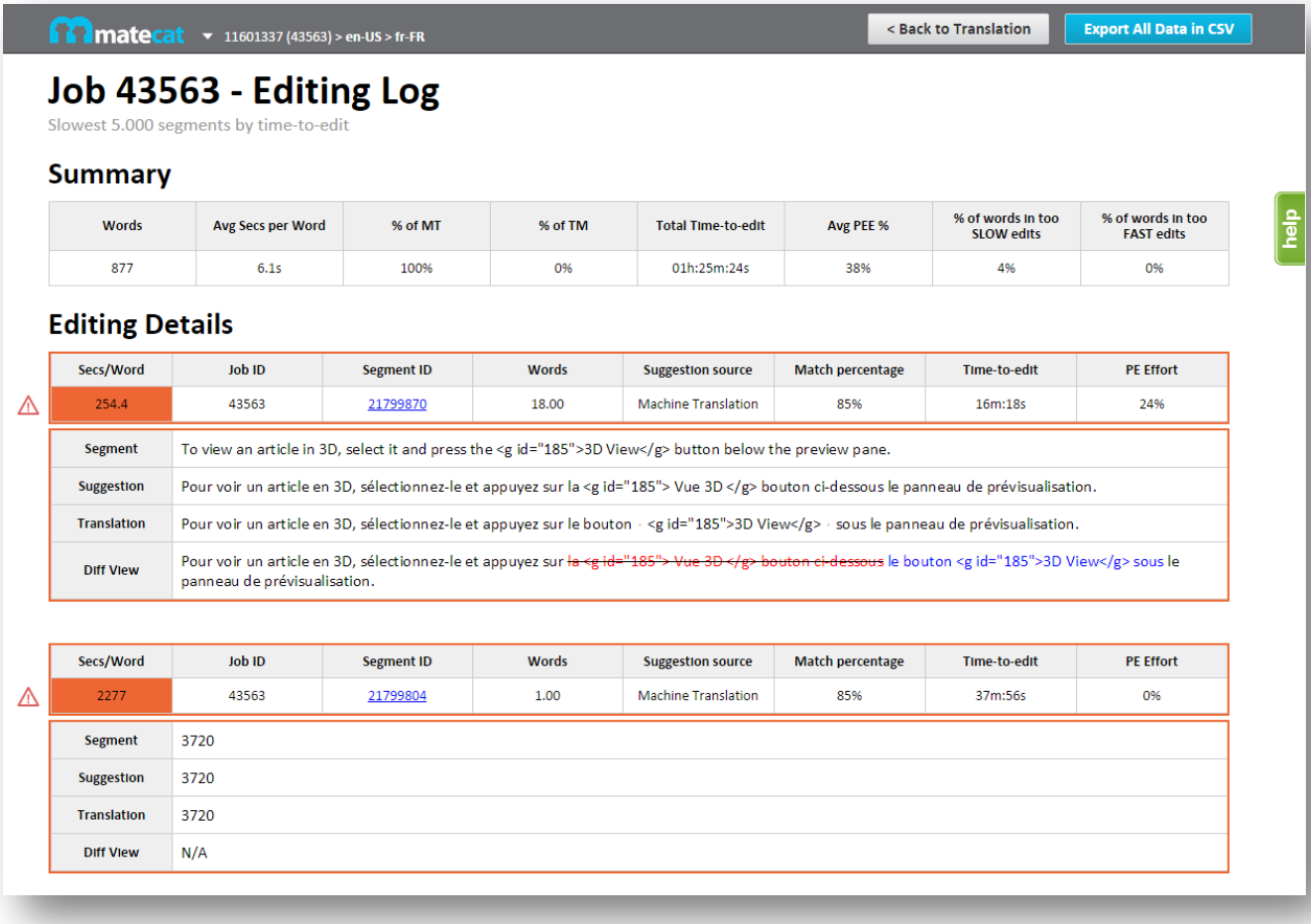
To enable **Safari spell-checker** for your target language, follow these steps:

- Go to Safari>Edit>Spelling and Grammar
- Show spelling and Grammar
- Choose your language in the dropdown menu
- Click on change

Editing Log

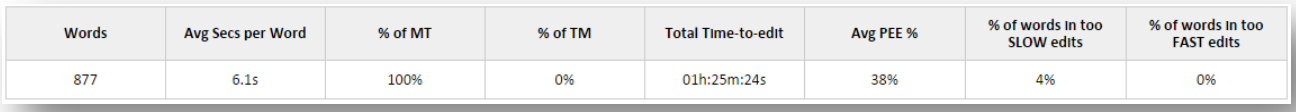
You can access the Editing Log from the corresponding button at the bottom of the Translation Editor page.

The Editing Log contains **statistical information about the translation**.


matecat
11601337 (43563) > en-US > fr-FR
< Back to Translation
Export All Data in CSV
Job 43563 - Editing Log
Slowest 5.000 segments by time-to-edit
Summary
Words | Avg Secs per Word | % of MT | % of TM | Total Time-to-edit | Avg PEE % | % of words in too SLOW edits | % of words in too FAST edits
877 | 6.1s | 100% | 0% | 01h:25m:24s | 38% | 4% | 0%
Editing Details
Secs/Word | Job ID | Segment ID | Words | Suggestion source | Match percentage | Time-to-edit | PE Effort
254.4 | 43563 | 21799870 | 18.00 | Machine Translation | 85% | 16m:18s | 24%
Segment | To view an article in 3D, select it and press the <g id="185">3D View</g> button below the preview pane.
Suggestion | Pour voir un article en 3D, sélectionnez-le et appuyez sur la <g id="185"> Vue 3D </g> bouton ci-dessous le panneau de prévisualisation.
Translation | Pour voir un article en 3D, sélectionnez-le et appuyez sur le bouton <g id="185">3D View</g> sous le panneau de prévisualisation.
Diff View | Pour voir un article en 3D, sélectionnez-le et appuyez sur la <g id="185"> Vue 3D </g> bouton ci-dessous le bouton <g id="185">3D View</g> sous le panneau de prévisualisation.
Secs/Word | Job ID | Segment ID | Words | Suggestion source | Match percentage | Time-to-edit | PE Effort
2277 | 43563 | 21799804 | 1.00 | Machine Translation | 85% | 37m:56s | 0%
Segment | 3720
Suggestion | 3720
Translation | 3720
Diff View | N/A
help

In particular:

- under the Summary section you can find statistical information about the entire project:

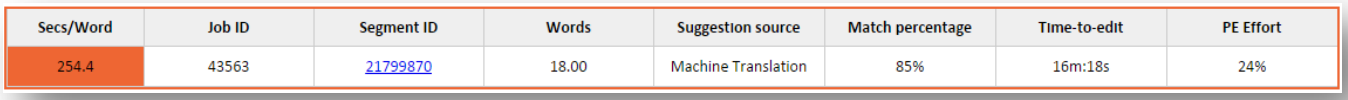


Words	Avg Secs per Word	% of MT	% of TM	Total Time-to-edit	Avg PEE %	% of words in too SLOW edits	% of words in too FAST edits
877	6.1s	100%	0%	01h:25m:24s	38%	4%	0%

Words	Number of **translated words**
Avg Secs per Word	**Average number of seconds** you needed to translate each word
% of MT	Percentage of **Machine Translation suggestions**
% of TM	Percentage of **translation memory suggestions**
Total Time-to-edit	**Total time** necessary to complete the translation so far
Avg PEE %	**Average post-editing effort** as a percentage. This indicates your efforts in updating machine translation suggestions

% of words in too SLOW edits	Percentage of words on which you spent **too much time** editing
% of words in too FAST edits	Percentage of words on which you spent **too little time** editing

Under the Editing details sections you can see all details per segment:



Secs/Word	Job ID	Segment ID	Words	Suggestion source	Match percentage	Time-to-edit	PE Effort
254.4	43563	21799870	18.00	Machine Translation	85%	16m:18s	24%

Secs/Word	**Seconds spent per word** in that segment
Job ID	MateCat **job ID**
Segment ID	**ID of the segment**. You can click on it to open the segment in the Translation Editor page.
Words	**Number of words** in the segment
Suggestion source	**Source** of the first suggestion
Match percentage	**Match percentage** of the first

	suggestion. It is always 85% by default for Machine Translation suggestions
Time-to-edit	**Time necessary** to edit the segment
PE Effort	**Post-editing effort** to edit the segment
Segment	Original **source** segment
Suggestion	First MateCat **suggestion**
Translation	**Final translation** approved by the translator
Diff View	**Differences** between first suggestion and final approved translation highlighted with track changes

The warning icon and orange borders indicate that you spent too much or too little time to edit the segment.

Finalising a Project

Once the translation is complete (100% on the progress bar), the **Preview** button at the top of the page is replaced by the **Download translation** button.

By clicking on the **Download Translation** button, you can download the translated text in the same format as the original file. If the project contains more than one file, all will be downloaded as compressed in a zip folder.

The text can be downloaded at any time during translation by clicking on the **Preview** button. The portions of text not yet translated will be replaced by a machine translation.

Managing Projects with MateCat

Management Panel

When you create a project, you have the option of adding the project to your **Management Panel** by logging in with your **Google account** before creating it.

This panel shows you the list of **all projects created using your Google account** and allows you to check the progress and status of your projects or recover them if needed.

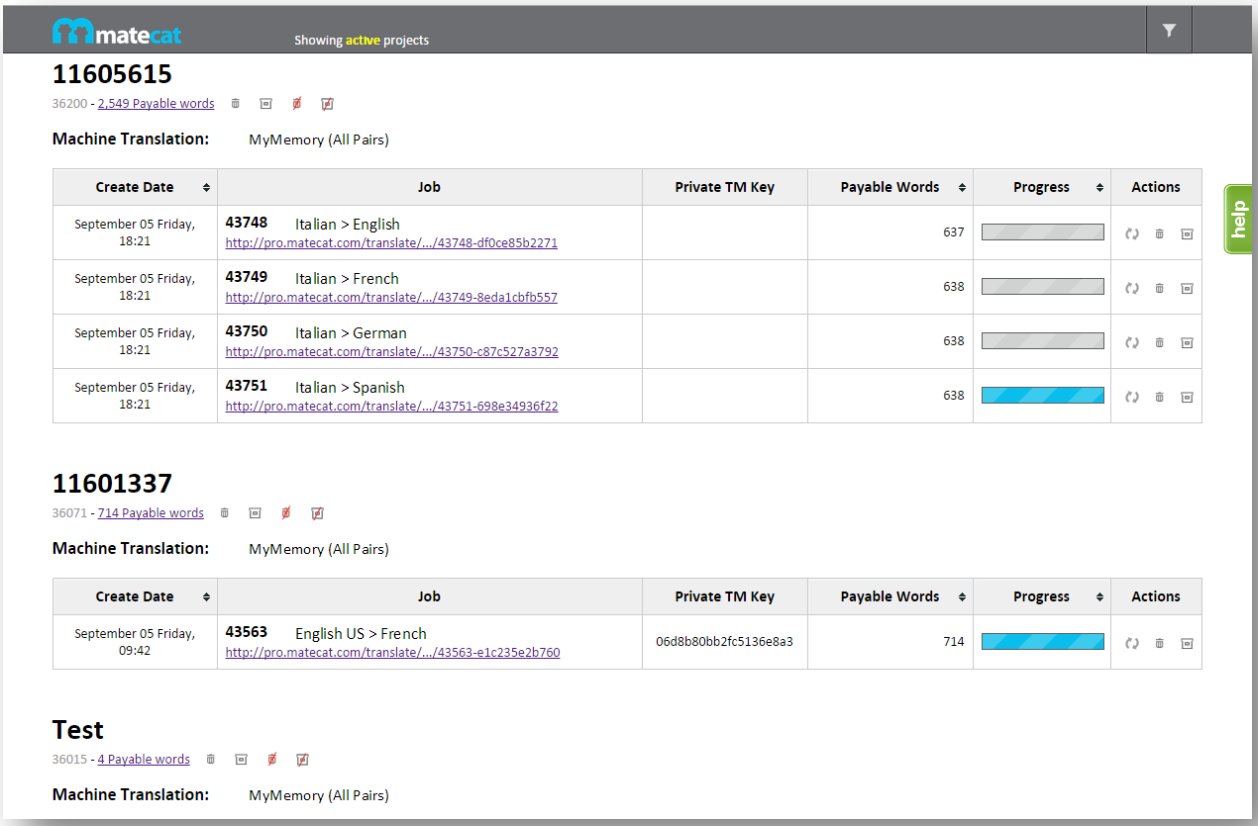
To log in with your Google account, click on the **Login** button at the bottom of the MateCat home page and sign in using your Google account.

> **Note**: remember to **always log in before creating a project** if you want to add it to your **Management Panel**. If you create a project **without logging in** with your Google account, the **job URL** of the project will not be saved automatically so please remember to save it manually in a safe place in order to be able to recover it easily if needed.

To access the **Management panel** you can:

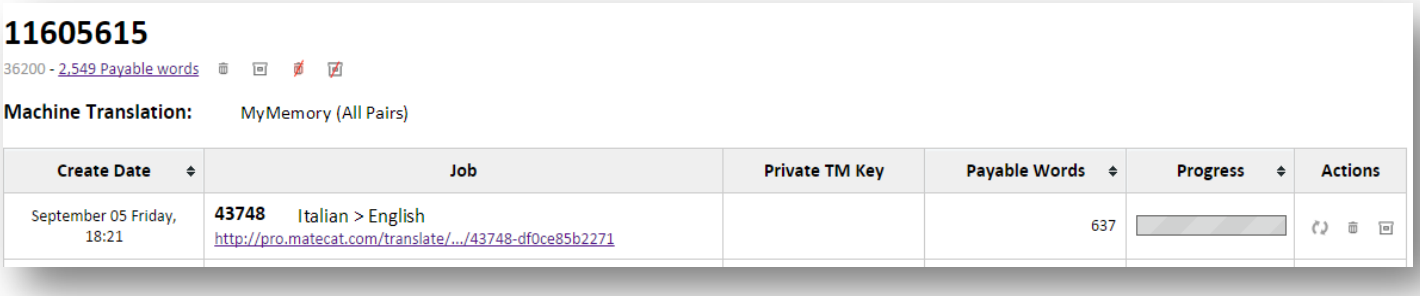
- Go to http://www.matecat.com/manage/1, or
- Click on Manage at the bottom of the MateCat home page

All relevant information is available for each project, as shown in the screenshot below:



Management panel

For each project you have:



- Project name (11605615)
- MateCat Project ID (36200)
- Total number of words in the project and link to the Volume Analysis Page (2,549 payable words)
- Machine Translation Server (MyMemory (All Pairs))
- Creation date (September 05 Friday, 18:21)
- Link to the translation jobs divided per language combination and MateCat job ID (under "Job" column) - you can access the Translation Editor window of each job by clicking on the corresponding job URL
- Private TM Key (if any; blank column if not present)
- Payable words per job (under Payable Words column)
- Progress bar of each translation job (under "Progress" column)
- A series of buttons that enable you to perform a number of tasks:

Change job password – This modifies the password of the job.

Cancel job/project – This deletes the job/project and disables the URL.

Archive Job/project – This prevents any changes from being applied to the job/project. However, the URL will still work and is accessible in r**ead-only mode**.

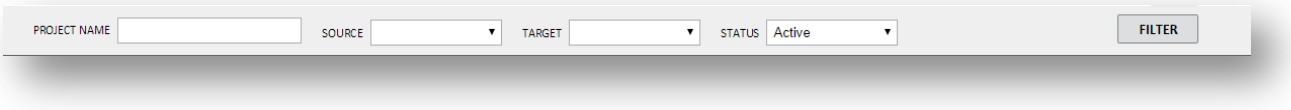
Resume project

Unarchive project

By default, the management panel will show your **active projects** only, in **chronological order**.

You can filter and see the cancelled and archived projects by clicking on the **Filter** button or on “Showing active projects” at the top of the page.

The following **search window** will be displayed:



You can filter your search by:

- Project name
- Source or target language (or both)
- Status (Active, Archived or Cancelled)

Click on the **FILTER** button to apply your criteria to the search.

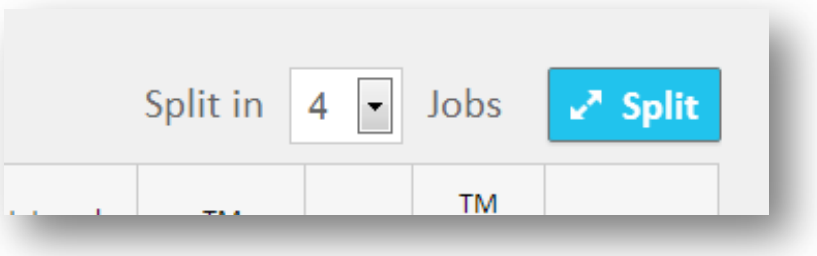
For reasons of privacy, **your projects are automatically archived after 30 days of inactivity**. You will still be able to reactivate them in your Management panel or in the Management panel of the person who created the project.

Splitting Jobs

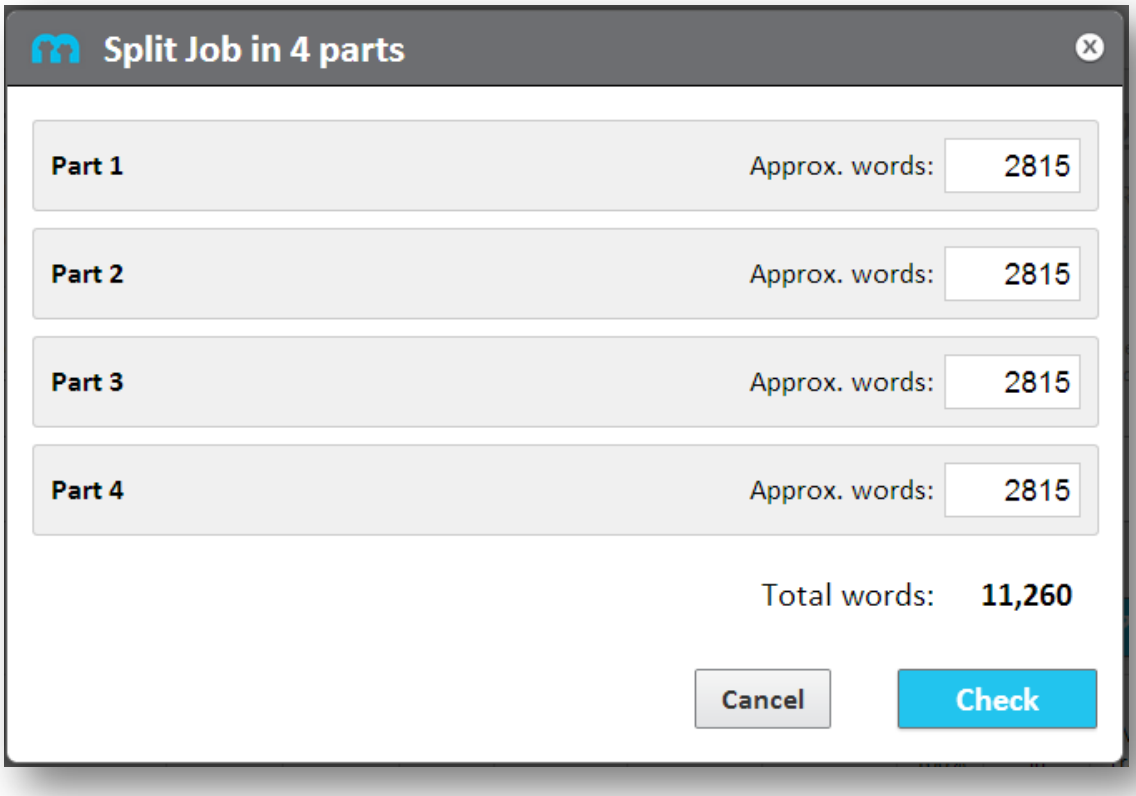
MateCat provides an easy-to-use split functionality **to divide large jobs into smaller parts and assign them to multiple translators**. This is useful for large **collaborative projects**.

You can split a job from the Volume Analysis page.

After completion of the volume analysis, you can select **how many jobs** you would like to create and click on the **Split** button. In the following example, we are going to create 4 jobs:

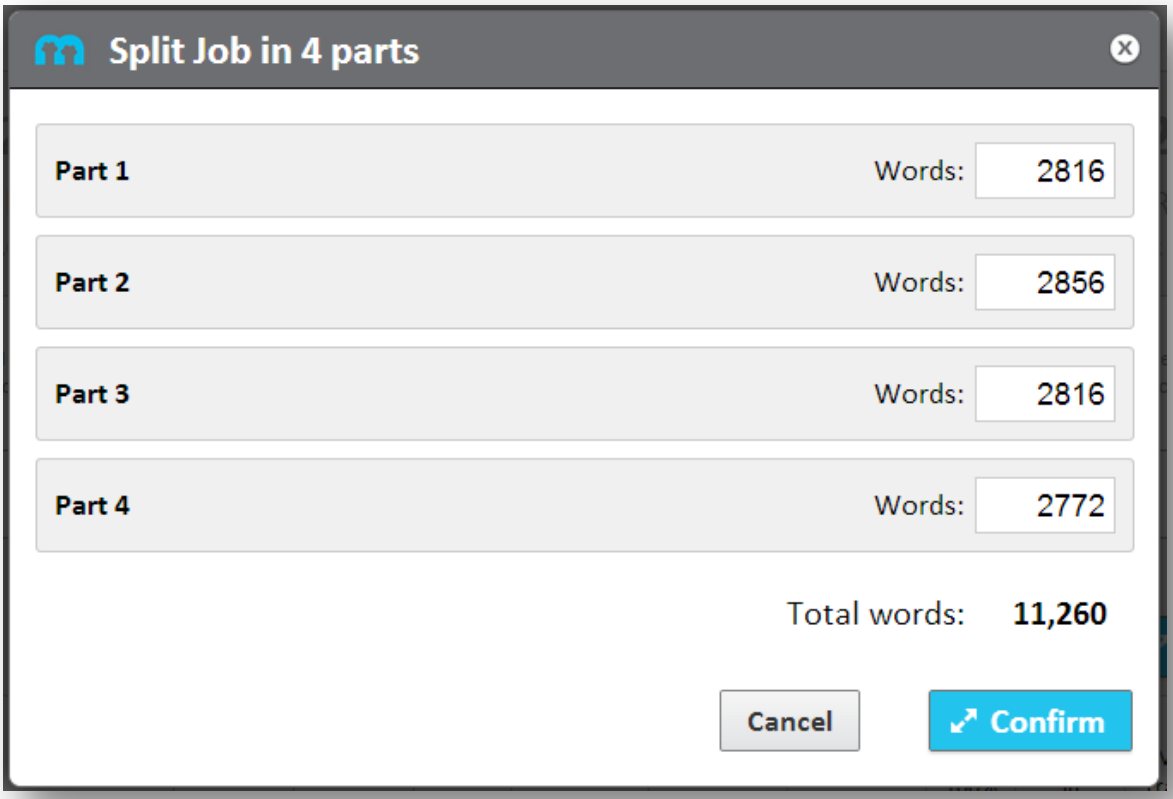


You will then be presented with a dialogue window where you can enter the approximate number of words for each job.



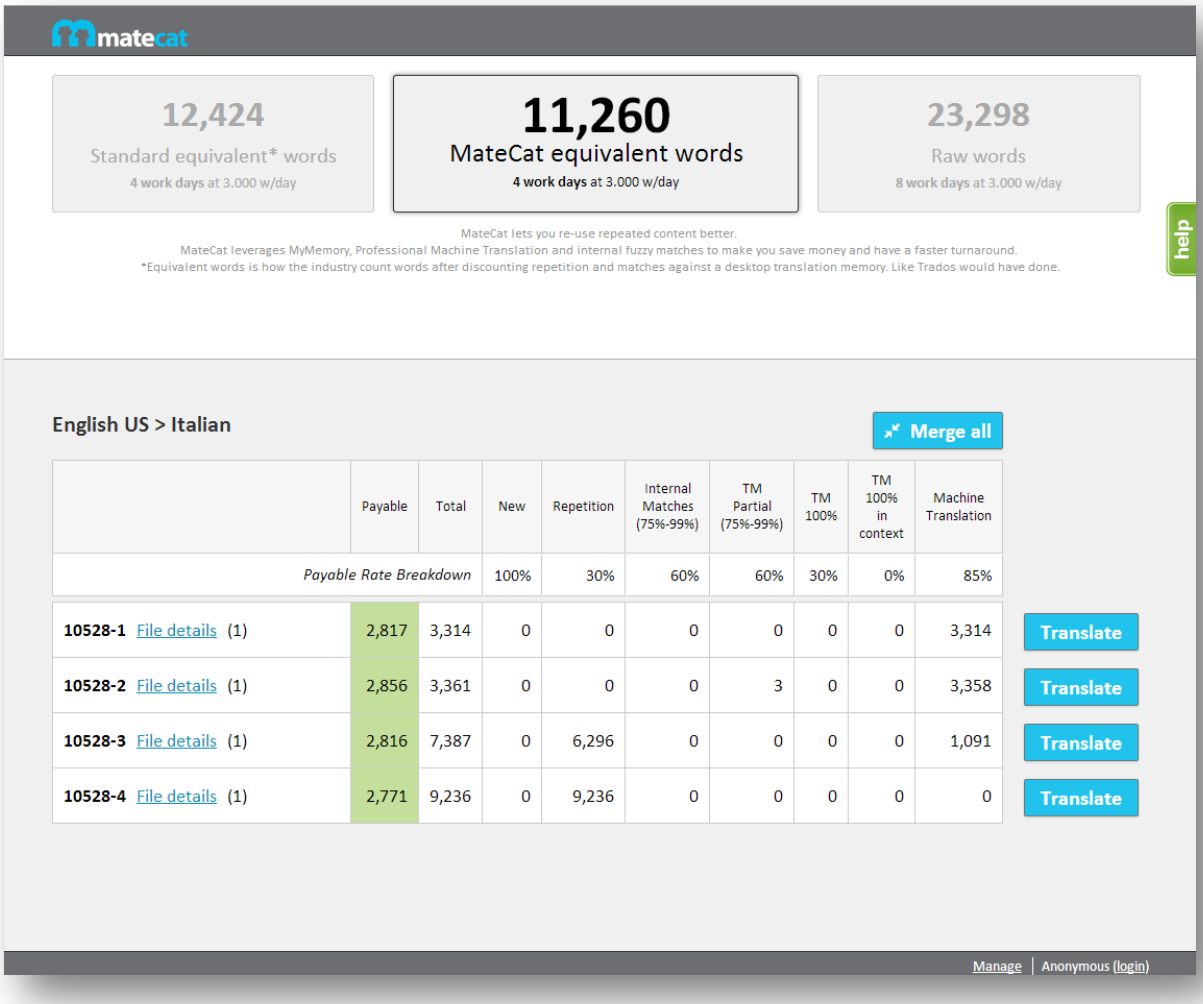
Splitting a job in 4 parts. The word count is approximate until you click on the Check button.

When you click on the **Check** button, MateCat checks each part and divides segments into an indicative number of words, avoiding any overlap between segments.



Splitting a job in 4 parts after checking the word count. Note that the word count for each part has changed.

When you click on **Confirm**, you will be presented with a **detailed analysis of the split jobs** and a **Translate** button for each part.



A job split in four parts

In the event that you make a mistake and would like to undo the splitting, just click on the **Merge all** button and the different parts will be merged back into the original job.

You just need to send each translator the **Job URL** for the part they will be working on.

When clicking on their link, translators will only see the part of the document assigned to them as editable. **They can also see the rest of the file, but they will not be able to edit it.** They will see the rest of the document being translated **in real time** and can refer to the other parts for comparison.

They will also **share a private TM key**, if previously associated, where all translations will be stored, so **each translator will also receive suggestions for segments already translated by the other translators** within the same project.

Each translation is complete once **all segments** of that job have been translated and marked as translated. The progress bar at the bottom of the screen should be blue and the To-do count should be zero.

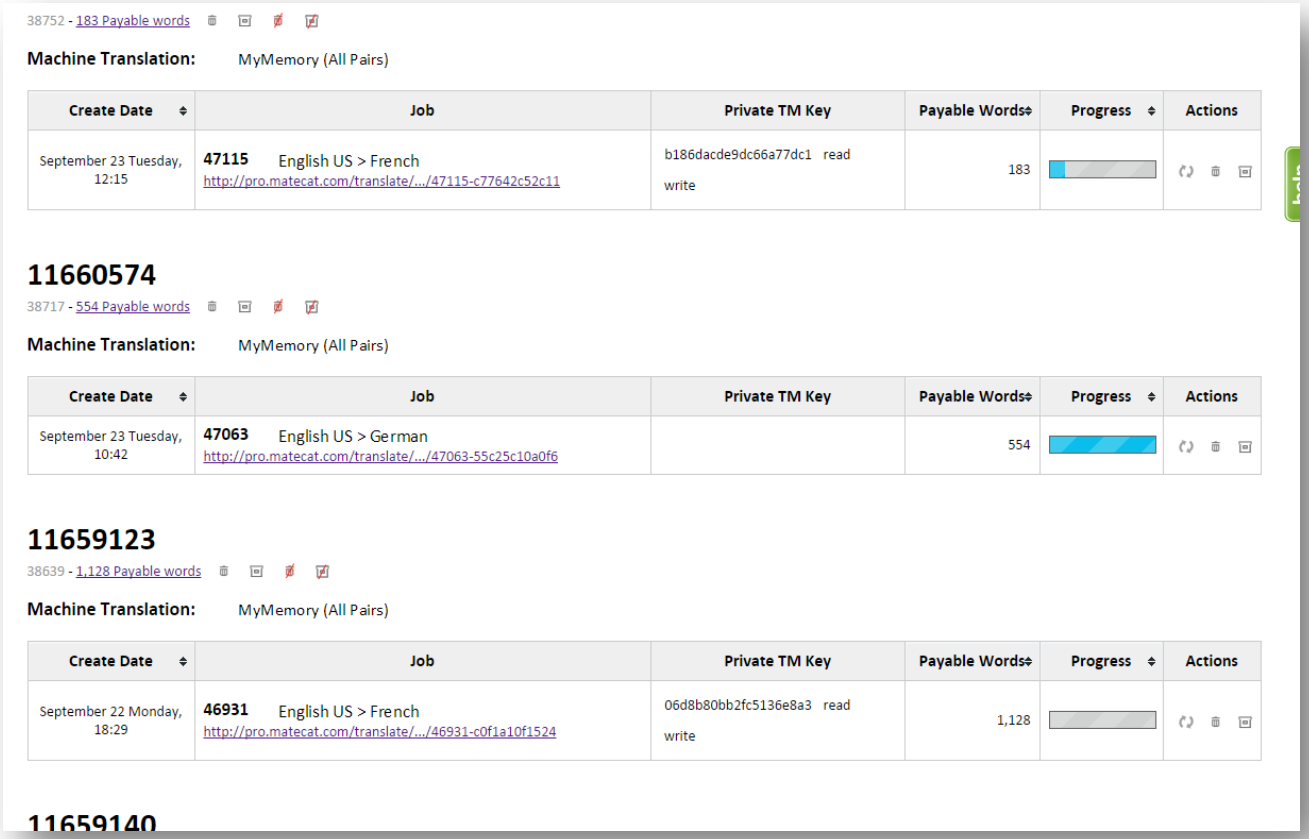
The progress bar, the To-do count and the payable words count only refer to the current job. You can check the status and progress of each job from your own Management Panel. Once all jobs are 100% complete, you can click on any of the links and click on the **Download translation** button. In this way you will download the **whole translated file**.

Resources for LSPs

MateCat can help LSPs (Language Service Providers) to manage projects easily and perform all checks not only after delivery by the translator but also **during the translation job** itself.

The main tools and functions available for Project Managers are the Management Panel, the Editing log and the QA messages in the Translation Editor window.

Through the Management Panel, the Project Manager can easily check **the status** of all translation jobs in **real time**. Here, he/she has a list of all active projects in chronological order and the progress bar for each job.



By clicking on the job URL of a specific project, it is possible to **check errors during the translation** and promptly advise the translator working on the project. Through the QA module, it is possible to check if there are **tag errors or translation conflicts** during the translation, without waiting for delivery by the translator and thus save time.

From the Translation Editor window, it is possible to access the Editing Log page by clicking on the **Editing Log** button at the bottom of the page.

In the Editing Log, the Project Manager can check the **time** spent by the translator for editing each segment, the total time spent by the translator to translate the entire file and check the **edits** made to suggestions. The changes are highlighted with track changes. This can help the Project Manager identify the Post-Editing efforts of the translator to edit the segments.

Appendix

Keyboard Shortcuts

MateCat relies heavily on **keyboard shortcuts** for standard and advanced functionalities. Getting to know the shortcuts below will help you be more productive when translating in MateCat.

Functionality	**Windows**	**Mac**
Confirm translation (click on Translated)	CTRL+Enter	CTRL+Enter CMD+Enter
Confirm translation and go to next untranslated segment (click on [T+>>])	CTRL+SHIFT+Enter	CTRL+SHIFT+Enter CMD+SHIFT+Enter
Go to next segment	CTRL+Down	CTRL+Down

Return to previous segment	CTRL+Up	CTRL+Up
Go to current segment	CTRL+Home	CMD+SHIFT+up
Copy source to target	ALT+CTRL+i	ALT+CTRL+i
Undo in segment (available in active segment)	CTRL+z	CTRL+z CMD+z
Redo in segment (available in active segment)	CTRL+y	CMD+SHIFT+z
Go to beginning of the line	Home	CMD+left
Go to end of the line	End	CMD+right
Move the cursor word by word to the right	CTRL+right	ALT+right
Move the cursor word by word to the left	CTRL+left	ALT+left
Open search (if not yet opened)	CTRL+f	CTRL+f CMD+f

Perform Concordance search on word(s) selected in the source or target segment	ALT+ c	ALT+ c
Use Translation suggestions (first/second/third suggestion)	CTRL+1/2/3	CTRL+1/2/3
Add a new line in the target segment	SHIFT+Enter	SHIFT+Enter
Add a non-breaking space	CTRL+Shift+Space	CTRL+Shift+Space

Installing an open source version of MateCat

We provide a guide meant for users who want to install and administer the open source version on their own machines.

It is available at www.matecat.com/installation-guide.